GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 13:04:25 ; Search time 935 Seconds
(without alignments)
12240.273 Million cell updates/sec

Title: US-09-497-822C-18
Perfect score: 5082
Sequence: 1 gagctctggacaaaattgag.....acaagcaaacaaaaaaaaaa 5082

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
	1	4188.2	82.4	4321	24	AAD30440	Human androgen rec
	2	3595.8	70.8	3715	20	AAZ23424	Human androgen rec
	3	3595.8	70.8	3715	24	ABN95802	Gene #2300 used to
	4	3594.2	70.7	3715	12	AAQ12001	Full-length human
	5	3547	69.8	3569	18	AAT63407	Androgen receptor
	6	3545.4	69.8	3569	10	AAN91772	Human androgen rec
c	7	3498.2	68.8	3590	23	AAS93472	DNA encoding novel
	8	2486.4	48.9	4180	10	AAN91773	Rat androgen recep
c	9	2244.2	44.2	3715	24	ABK34013	Human DNA for stag
c	10	2244.2	44.2	3715	24	ABK31545	Signal transductio
	11	2189.2	43.1	3217	12	AAQ12002	Full-length rat an
	12	2185.8	43.0	3217	10	AAN91578	Rat androgen recep
	13	2154	42.4	3715	24	ABK34012	Human DNA for stag
	14	2154	42.4	3715	24	ABK31544	Signal transductio
	15	1715.6	33.8	1810	22	AAF84342	Human androgen rec
c	16	1587	31.2	1998	23	AAS91973	DNA encoding novel
	17	1532.6	30.2	6905	24	ABK29886	Andorgen receptor
	18	1340	26.4	2517	24	ABA01683	Fused androgen rec
	19	1148	22.6	1329	24	ABN85656	Human androgen rec
	20	834.4	16.4	1171	24	ABN85657	Human androgen rec
c	21	761.6	15.0	1432	24	ABL34259	Human immune syste
	22	711.8	14.0	1893	12	AAQ12008	TrpE/androgen rece
	23	703.6	13.8	1731	12	AAQ12007	TrpE/N-terminal do
	24	703.2	13.8	1432	24	ABL34258	Human immune syste
	25	611.6	12.0	612	21	AAC70311	Single nucleotide
	26	611.6	12.0	612	21	AAC70335	Single nucleotide
	27	575.6	11.3	598	21	AAC70305	Single nucleotide
	28	575.6	11.3	598	21	AAC70308	Single nucleotide
	29	524.2	10.3	5168	24	ABK94716	Neurodegenerative
	30	401.6	7.9	6244	16	AAQ76269	PhCMV*-1-controlle
	31	401.6	7.9	6244	20	AAX81724	cDNA encoding rabb
	32	401.6	7.9	6244	20	AAX27905	Rabbit progesteron
	33	401.6	7.9	6244	20	AAX01369	PhCMV*-1 rabbit pr
	34	401.6	7.9	6244	22	AAD09837	Rabbit progesteron
	35	400	7.9	6244	16	AAT06872	PhCMV*-1 promoter
	36	384	7.6	1231	14	AAQ37761	Sequence of a 1.23
	37	384	7.6	3014	18	AAT94756	Human progesterone
	38	384	7.6	3014	18	AAT84747	Human progesterone
	39	383	7.5	2970	21	AAA53851	Sequence encoding
	40	343.2	6.8	360	22	AAL19609	Human breast cance
	41	340.6	6.7	1329	12	AAQ12009	TrpE/AR androgen-b
	42	320.8	6.3	358	21	AAC08780	Human secreted pro
	43	310.8	6.1	6322	24	ABK63639	Rat sequence diffe
	44	310.8	6.1	7257	18	AAT68656	Green fluorescent
	45	305.6	6.0	5749	24	ABL61813	Colon adenocarcino

ALIGNMENTS

RESULT 1
AAD30440
ID AAD30440 standard; DNA; 4321 BP.
XX
AC AAD30440;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human androgen receptor (AR) gene.
XX
KW Human; AIB1; amplified in breast cancer 1; androgen receptor; AR;
KW prostate cancer; chromosome X; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1115..3874
FT /*tag= a
FT /product= "Human androgen receptor (AR) protein"
XX
PN WO200210452-A2.
XX
PD 07-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-US23834.
XX
PR 27-JUL-2000; 2000US-221074P.
XX
PA (UYRP) UNIV ROCHESTER.
XX
PI Chang C;
XX

DR WPI; 2002-206195/26.
DR P-PSDB; AAE19061.
XX
PT Assessing the risk of acquiring or developing prostate cancer in a
PT human subject, comprises determining the length of the contiguous CAG,
PT CAA and/or GGN repeats in the AIB1 gene and/or androgen receptor gene
PT of the subject -
XX
PS Disclosure; Page 61-62; 86pp; English.
XX
CC The invention relates to a method for assessing the risk of prostate
CC cancer in a human subject. The method involves determining the length
CC of the contiguous CAG or CAA repeats in both AIB1 (Amplified In Breast
CC cancer 1) gene alleles or contiguous CAG, CAA or GGN repeats in the
CC androgen receptor gene of the subject. The method is useful for
CC assessing a subject's risk for acquiring or developing prostate cancer.
CC The present sequence is human androgen receptor (AR) gene. Human
CC AR gene is located on X chromosome.
XX
SQ Sequence 4321 BP; 966 A; 1281 C; 1168 G; 906 T; 0 other;

Query Match 82.4%; Score 4188.2; DB 24; Length 4321;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 4297; Conservative 0; Mismatches 16; Indels 24; Gaps 8;

```
Qy    754 CGAGATCCCGGGGAGCCAGCTTGCTGGGAGAGCGGGAACGGTCCGGAGCAAGCCCAGAGG 813
          ||||||||||||||||||||||||||||||||||| |||||||||||||||||||| |||
Db      1 CGAGATCCCGGGGAGCCAGCTTGCTGGGAGAGCGGG-ACGGTCCGGAGCAAGCCCACAGG 59

Qy    814 CAGAGGAGGCGACAGAGGGAAAAAGGGCCCNAGCTAGCCGCTCCAGTGCTGTACAGNAGC 873
          |||||||||||||||||||||||||||| || |||||||||||||||||||||||| |||
Db     60 CAGAGGAGGCGACAGAGGGAAAAAGGG-CCGAGCTAGCCGCTCCAGTGCTGTACAGGAGC 118

Qy    874 CGAA-GGACGCACCACGCCAGCCCCAGCCCGGCTCCAGCGACAGCNAACGCCTCTTGCA- 931
          |||| |||||||||||||||||||||||||||||||||||||||||| ||||||||||| 
Db    119 CGAAGGGACGCACCACGCCAGCCCCAGCCCGGCTCCAGCGACAGCCAACGCCTCTTGCAG 178

Qy    932 -----NGCGTTCGAAGCCGCCGCCCGGAGCTGCCCTTTCCTCTTCGGTGAAGTTTTTAAA 986
               | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    179 CGCGGCGGCTTCGAAGCCGCCGCCCGGAGCTGCCCTTTCCTCTTCGGTGAAGTTTTTAAA 238

Qy    987 AGCTGCTAAAGACTCGGAGGAAGCAAGGAAAGTGCCTGGTAGGACTGACGGCTGCCTTTG 1046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    239 AGCTGCTAAAGACTCGGAGGAAGCAAGGAAAGTGCCTGGTAGGACTGACGGCTGCCTTTG 298

Qy   1047 TCCTCCTCCTCTCCACCCCGCCTCCCCCCACCCTGCCTTCCCCCCCTCCCCCGTCTTCTC 1106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 TCCTCCTCCTCTCCACCCCGCCTCCCCCCACCCTGCCTTCCCCCCCTCCCCCGTCTTCTC 358

Qy   1107 TCCCGCAGCTGCCTCAGTCGGCTACTCTCAGCCAACCCCCCTCACCACCCTTCTCCCCAC 1166
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 TCCCGCAGCTGCCTCAGTCGGCTACTCTCAGCCAACCCCCCTCACCACCCTTCTCCCCAC 418

Qy   1167 CCGCCCCCCCCGCCCCCGTCGGCCCAGCGNTGNCAGNCCGAGTTTGCAGAGAGGTAACTCC 1226
          |||||||||||||||||||||||||||| || ||| ||||||||||||||||||||||||
Db    419 CCGCCCCCCCCGCCCCCGTCGGCCCAGCGCTGCCAGCCCGAGTTTGCAGAGAGGTAACTCC 478

Qy   1227 CTTTGGCTGCGAGCGGGCGAGNCTAGCTGCACATTGCAAAGAAGGCTCTTAGGAG-CAGG 1285
          |||||||||||||||||||| |||||||||||||||||||||||||||||||| ||||
Db    479 CTTTGGCTGCGAGCGGGCGAG-CTAGCTGCACATTGCAAAGAAGGCTCTTAGGAGCCAGG 537

Qy   1286 CGACTGGGGAGCGGCTTCAGCACTGCAGCCACGACCNGCCTGGTTAGGCTGCACGCGGAG 1345
          |||||||||||||||||||||||||||||||||||| |||||||||  | |  ||||||
Db    538 CGACTGGGGAGCGGCTTCAGCACTGCAGCCACGACCCGCCTGGTTAGAATTCCGGCGGAG 597

Qy   1346 AGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCCCACCCCGAGTG 1405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 AGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCCCACCCCGAGTG 657

Qy   1406 CGGAGCCAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCAAAAAACAAAAC 1465
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    658 CGGAG-CAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCAAAAAACAAAAC 716

Qy   1466 AAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTCAGTTCTTATTTGCAC 1525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    717 AAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTCAGTTCTTATTTGCAC 776

Qy   1526 CTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTCTTTTAAGATCT 1585
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    777 CTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTCTTTTAAGATCT 836

Qy   1586 GGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGCCTAGCAGGGCA 1645
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    837 GGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGCCTAGCAGGGCA 896

Qy   1646 GATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGGCTGTCAGAGCGCTTTTT 1705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    897 GATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGGCTGTCAGAGCGCTTTTT 956

Qy   1706 GCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGGCAGCTAGCTGC 1765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    957 GCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGGCAGCTAGCTGC 1016

Qy   1766 AGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAAGGGGAGGCGGG 1825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1017 AGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAAGGGGAGGCGGG 1076

Qy   1826 GTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCAGTTAGGGCTGG 1885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1077 GTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCAGTTAGGGCTGG 1136

Qy   1886 GAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCAGAATCTGTTCC 1945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1137 GAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCAGAATCTGTTCC 1196

Qy   1946 AGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGCCGCGAGCGCAG 2005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1197 AGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGCCGCGAGCGCAG 1256

Qy   2006 CACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2065
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1257 CACCTCCCGGCGCCAGTTTGCTGCTGCT-----------GCAGCAGCAGCAGCAGCAGC 1304

Qy   2066 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGACTAGCCCCAGGC 2125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1305 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGACTAGCCCCAGGC 1364

Qy   2126 AGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAGAGGCCCCACAG 2185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1365 AGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAGAGGCCCCACAG 1424

Qy   2186 GCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGCCCTGGAGTGCC 2245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1425 GCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGCCCTGGAGTGCC 1484

Qy   2246 ACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAGCAAGGGGCTGC 2305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1485 ACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAGCAAGGGGCTGC 1544

Qy   2306 CGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATCCACGTTGTCCC 2365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1545 CGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATCCACGTTGTCCC 1604

Qy   2366 TGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAAAGACATCCTGA 2425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1605 TGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAAAGACATCCTGA 1664

Qy   2426 GCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGTATCCGAAGGCA 2485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1665 GCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGTATCCGAAGGCA 1724

Qy   2486 GCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAAGGACAATTACT 2545
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1725 GCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAAGGACAATTACT 1784

Qy   2546 TAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGCAGTGTCGGTGT 2605
```

```
Db  1785 TAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGCAGTGTCGGTGT 1844
Qy  2606 CCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACAGCTTCGGGGGG 2665
Db  1845 CCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACAGCTTCGGGGGG 1904
Qy  2666 ATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCCTTGTGCCC 2725
Db  1905 ATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCCTTGTGCCC 1964
Qy  2726 CATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAGCACTGAAGATA 2785
Db  1965 CATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAGCACTGAAGATA 2024
Qy  2786 CTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGGCGAGAGCCTAG 2845
Db  2025 CTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGGCGAGAGCCTAG 2084
Qy  2846 GCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCCGTCTACCCTGT 2905
Db  2085 GCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCCGTCTACCCTGT 2144
Qy  2906 CTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACTACTACA 2965
Db  2145 CTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACTACTACA 2204
Qy  2966 ACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCCACGCTC 3025
Db  2205 ACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCCACGCTC 2264
Qy  3026 GCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGGGGCGGCGCAGT 3085
Db  2265 GCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGCGGCGGCGCAGT 2324
Qy  3086 GCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTTCTGGGT 3145
Db  2325 GCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTTCTGGGT 2384
Qy  3146 CACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAGGCCAGT 3205
Db  2385 CACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAGGCCAGT 2444
Qy  3206 TGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGCG 3265
Db  2445 TGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGCG 2504
Qy  3266 GCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAGCTGTAGCCCCCTACGGCTACACTC 3325
Db  2505 GCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAGCTGTAGCCCCCTACGGCTACACTC 2564
Qy  3326 GGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCACCGCACCTGATGTGTGGTACC 3385
Db  2565 GGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCACCGCACCTGATGTGTGGTACC 2624
Qy  3386 CTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCACTTGTGTCAAAAGCGAAATGG 3445
Db  2625 CTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCACTTGTGTCAAAAGCGAAATGG 2684
Qy  3446 GCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACATGCGTTTGGAGACTGCCAGGG 3505
Db  2685 GCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACATGCGTTTGGAGACTGCCAGGG 2744
Qy  3506 ACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAAGACCTGCCTGATCTGTGGAG 3565
Db  2745 ACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAAGACCTGCCTGATCTGTGGAG 2804
Qy  3566 ATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACATGTGGAAGCTGCAAGGTCTTCTTCA 3625
Db  2805 ATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACATGTGGAAGCTGCAAGGTCTTCTTCA 2864
Qy  3626 AAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAGCAGAAATGATTGCACTATTG 3685
Db  2865 AAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAGCAGAAATGATTGCACTATTG 2924
Qy  3686 ATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCGGAAATGTTATGAAGCAGGGA 3745
Db  2925 ATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCGGAAATGTTATGAAGCAGGGA 2984
Qy  3746 TGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCTGAAACTACAGGAGGAAGGAG 3805
Db  2985 TGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCTGAAACTACAGGAGGAAGGAG 3044
Qy  3806 AGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCAGAAGCTGACAGTGTCACACA 3865
Db  3045 AGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCAGAAGCTGACAGTGTCACACA 3104
Qy  3866 TTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATGTCCTGGAAGCCATTGAGCCAGGTG 3925
Db  3105 TTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATGTCCTGGAAGCCATTGAGCCAGGTG 3164
Qy  3926 TAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTTTGCAGCCTTGCTCTCTAGCC 3985
Db  3165 TAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTTTGCAGCCTTGCTCTCTAGCC 3224
Qy  3986 TCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAAGTGGGCCAAGGCCTTGCCTG 4045
Db  3225 TCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAAGTGGGCCAAGGCCTTGCCTG 3284
Qy  4046 GCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGGC 4105
Db  3285 GCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGGC 3344
Qy  4106 TCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGTCAACTCCAGGATGCTCTACT 4165
Db  3345 TCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGTCAACTCCAGGATGCTCTACT 3404
Qy  4166 TCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCCGGATGTACAGCCAGT 4225
Db  3405 TCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCCGGATGTACAGCCAGT 3464
Qy  4226 GTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCAAATCACCCCCCAGGAAT 4285
Db  3465 GTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCAAATCACCCCCCAGGAAT 3524
Qy  4286 TCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAATC 4345
Db  3525 TCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAATC 3584
Qy  4346 AAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTCGATCGTATCATTGCAT 4405
Db  3585 AAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTCGATCGTATCATTGCAT 3644
Qy  4406 GCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTACCAGCTCACCAAGCTCCTGG 4465
Db  3645 GCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTACCAGCTCACCAAGCTCCTGG 3704
Qy  4466 ACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCACTTTTGACCTGCTAATCAAGT 4525
Db  3705 ACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCACTTTTGACCTGCTAATCAAGT 3764
Qy  4526 CACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGAGATCATCTCTGTGCAAGTGC 4585
Db  3765 CACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGAGATCATCTCTGTGCAAGTGC 3824
Qy  4586 CCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAA 4645
Db  3825 CCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAA 3884
Qy  4646 ACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTGTTATAACTC 4705
Db  3885 ACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTGTTATAACTC 3944
Qy  4706 TGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTATTGATGTACAGTCTGTCATGA 4765
Db  3945 TGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTATTGATGTACAGTCTGTCATGA 4004
```

```
Qy   4766 ACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTCTCTTTCTCTCCTTTCTTTTTCT 4825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4005 ACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTCTCTTTCTCTCCTTTCTTTTTCT 4064

Qy   4826 TCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTC 4885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4065 TCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTC 4124

Qy   4886 CTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCC 4945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4125 CTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCC 4184

Qy   4946 AGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCAGCCACACAAACGTTTACTTA 5005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4185 AGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCAGCCACACAAACGTTTACTTA 4244

Qy   5006 TCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAATCAAAACAAAAAACA 5065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4245 TCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAATCAAAACAAAAAACA 4304

Qy   5066 AGCAAACAAAAAAAAAA 5082
          ||||||||||||||||||
Db   4305 AGCAAACAAAAAAAAAA 4321
```

RESULT 2
AAZ23424
ID AAZ23424 standard; DNA; 3715 BP.
XX
AC AAZ23424;
XX
DT 19-JAN-2000 (first entry)
XX
DE Human androgen receptor DNA.
XX
KW Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;
KW huntingtin polypeptide; Machado-Joseph disease; SCA1; SCA2; SCA6;
KW atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;
KW Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;
KW dentatorubropallidoluysian atrophy; cell proliferation; cell survival;
KW neoplastic; malignant; autoimmune; fibrotic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 532..3288
FT /*tag= a
FT /product= "androgen receptor"
XX
PN WO9945944-A1.
XX
PD 16-SEP-1999.
XX
PF 11-MAR-1999; 99WO-US05250.
XX
PR 12-MAR-1998; 98US-0041886.
XX
PA (BURN-) BURNHAM INST.
XX
PI Bredesen DE, Rabizadeh S;
XX
DR WPI; 1999-561617/47.
DR P-PSDB; AAY33491.
XX
PT New proapoptotic dependence peptides, used to develop products for
PT treating, e.g. Alzheimer's disease -
XX
PS Disclosure; Page 86-90; 199pp; English.
XX
CC This invention describes novel pure proapoptotic dependence peptides
CC which comprise a sequence of an active dependence domain selected from
CC dependence polypeptides consisting of p75NTR, androgen receptor, DCC,
CC huntingtin polypeptide, Machado-Joseph disease gene product, SCA1, SCA2,
CC SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable
CC of inducing cell death and can be used to develop products to mediate or
CC inhibit apoptosis. The methods can be used for reducing the severity of
CC a proapoptotic dependence domain mediated pathological conditions e.g.
CC Huntington's disease, Alzheimer's disease, Kennedy's disease,
CC Spinocerebellar ataxias, dentatorubropallidoluysian atrophy,
CC Machado-Joseph disease, stroke or head trauma. They can also be used for
CC reducing the severity of a pathological condition mediated by upregulated
CC cell proliferation or cell survival e.g. neoplastic, malignant,
CC autoimmune or fibrotic conditions. This sequence encodes a human
CC androgen receptor described in the method of the invention.
XX
SQ Sequence 3715 BP; 841 A; 1055 C; 1001 G; 818 T; 0 other;

Query Match 70.8%; Score 3595.8; DB 20; Length 3715;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3685; Conservative 0; Mismatches 7; Indels 45; Gaps 4;

```
Qy   1335 TGCACGCGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCC 1394
          | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4 TTCCGGCGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCC 63

Qy   1395 CACCCCGAGTGCGGAGCCAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCA 1454
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     64 CACCCCGAGTGCGGAG-CAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCA 122

Qy   1455 AAAAACAAAACAAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTCAGTT 1514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 AAAAACAAAACAAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTCAGTT 182

Qy   1515 CTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTC 1574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 CTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTC 242

Qy   1575 TTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGC 1634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 TTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGC 302

Qy   1635 CTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGGCTGTCA 1694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 CTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGGCTGTCA 362

Qy   1695 GAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGG 1754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 GAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGG 422

Qy   1755 CAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAA 1814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 CAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAA 482

Qy   1815 GGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCA 1874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 GGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCA 542

Qy   1875 GTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCA 1934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 GTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCA 602

Qy   1935 GAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGC 1994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 GAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGC 662

Qy   1995 CGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGCAGCAGCAGCAGCA 2054
          ||||||||||||||||||||||||||||||||||||||||
Db    663 CGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCT---------------------- 701

Qy   2055 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGAC 2114
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 ---GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGAC 758

Qy   2115 TAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAG 2174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    759 TAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAG 818

Qy   2175 AGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGC 2234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    819 AGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGC 878

Qy   2235 CCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAG 2294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    879 CCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAG 938

Qy   2295 CAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATC 2354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    939 CAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATC 998

Qy   2355 CACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAA 2414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    999 CACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAA 1058

Qy   2415 AGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGT 2474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1059 AGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGT 1118

Qy   2475 ATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAA 2534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1119 ATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAA 1178

Qy   2535 GGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGC 2594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1179 GGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGC 1238

Qy   2595 AGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACA 2654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1239 AGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACA 1298

Qy   2655 GCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCAC 2714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1299 GCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCAC 1358

Qy   2715 TCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAG 2774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1359 TCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAG 1418

Qy   2775 CACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGG 2834
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1419 CACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGG 1478

Qy   2835 CGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCC 2894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1479 CGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCC 1538

Qy   2895 GTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCG 2954
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1539 GTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCG 1598

Qy   2955 CGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCA 3014
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1599 CGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCA 1658

Qy   3015 TCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGC 3074
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1659 TCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGC 1718

Qy   3075 GGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACC 3134
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1719 GGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACC 1778

Qy   3135 CGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGA 3194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1779 CGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGA 1838

Qy   3195 AGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGT---------GGCGGCGG 3245
          ||||||||||||||||||||||||||||||||||||||||||||         ||||||||
Db   1839 AGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGTGGCGGCGGCGGCGG 1898

Qy   3246 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAGCTGT 3305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1899 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGAAGCTGT 1958

Qy   3306 AGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCAC 3365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1959 AGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCAC 2018

Qy   3366 CGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCAC 3425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2019 CGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCAC 2078

Qy   3426 TTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACAT 3485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2079 TTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACAT 2138

Qy   3486 GCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAA 3545
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2139 GCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAA 2198

Qy   3546 GACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACATGTGG 3605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2199 GACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACATGTGG 2258

Qy   3606 AAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAG 3665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2259 AAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAG 2318

Qy   3666 CAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCG 3725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2319 CAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCG 2378

Qy   3726 GAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCT 3785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2379 GAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCT 2438

Qy   3786 GAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCA 3845
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2439 GAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCA 2498

Qy   3846 GAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAAGTCCT 3905
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2499 GAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATGTCCT 2558

Qy   3906 GGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTT 3965
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2559 GGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTT 2618

Qy   3966 TGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAA 4025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2619 TGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAA 2678

Qy   4026 GTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCAT 4085
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2679 GTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCAT 2738

Qy   4086 TCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGT 4145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2739 TCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGT 2798

Qy   4146 CAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAA 4205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2799 CAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAA 2858

Qy   4206 GTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCT 4265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2859 GTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCT 2918

Qy   4266 CCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCC 4325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2919 CCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCC 2978
```

```
Qy   4326 AGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGA 4385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2979 AGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGA 3038

Qy   4386 ACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTA 4445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3039 ACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTA 3098

Qy   4446 CCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCAC 4505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3099 CCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCAC 3158

Qy   4506 TTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGA 4565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3159 TTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGA 3218

Qy   4566 GATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCA 4625
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3219 GATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCA 3278

Qy   4626 CACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATG 4685
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3279 CACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATG 3338

Qy   4686 TCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTAT 4745
          ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db   3339 TCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGTTTAATTTCCTCTAT 3398

Qy   4746 TGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTCTC 4805
          ||||||||||||||||||          ||||||||||||||||||||||||||||||||
Db   3399 TGATGTACAGTCTGTCATG----------GAATTCTATTTGCTGGGCTTTTTTTTTCTC 3447

Qy   4806 TTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGAC 4865
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3448 TTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGAC 3507

Qy   4866 TTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTG 4925
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3508 TTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTG 3567

Qy   4926 TGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCA 4985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3568 TGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCA 3627

Qy   4986 GCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGG 5045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3628 GCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGG 3687

Qy   5046 GGAAATCAAAACAAAAA 5062
          |||||||||||||||||
Db   3688 GGAAATCAAAACAAAAA 3704
```

RESULT 3
ABN95802
ID ABN95802 standard; DNA; 3715 BP.
XX
AC ABN95802;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2300 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample
XX
PS Claim 1; SEQ ID NO 2300; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3715 BP; 841 A; 1055 C; 1001 G; 818 T; 0 other;

Query Match 70.8%; Score 3595.8; DB 24; Length 3715;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3685; Conservative 0; Mismatches 7; Indels 45; Gaps 4;

```
Qy   1335 TGCACGCGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCC 1394
          | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4 TTCCGGCGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCC 63

Qy   1395 CACCCCGAGTGCGGAGCCAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCA 1454
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     64 CACCCCGAGTGCGGAG-CAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCA 122

Qy   1455 AAAAACAAAACAAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTCAGTT 1514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 AAAAACAAAACAAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTCAGTT 182

Qy   1515 CTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTC 1574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 CTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTC 242

Qy   1575 TTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGC 1634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 TTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGC 302

Qy   1635 CTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGGCTGTCA 1694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 CTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGGCTGTCA 362

Qy   1695 GAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGG 1754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 GAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGG 422

Qy   1755 CAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAA 1814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 CAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAA 482

Qy   1815 GGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCA 1874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    483 GGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCA 542

Qy   1875 GTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCA 1934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 GTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCA 602

Qy   1935 GAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGC 1994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 GAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGC 662

Qy   1995 CGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGCAGCAGCAGCAGCA 2054
          ||||||||||||||||||||||||||||||||||||
Db    663 CGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCT---------------------- 701

Qy   2055 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGAC 2114
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 ---GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGAC 758

Qy   2115 TAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAG 2174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    759 TAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAG 818

Qy   2175 AGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGC 2234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    819 AGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGC 878

Qy   2235 CCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAG 2294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    879 CCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAG 938

Qy   2295 CAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATC 2354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    939 CAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATC 998

Qy   2355 CACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAA 2414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    999 CACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAA 1058

Qy   2415 AGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGT 2474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1059 AGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGT 1118

Qy   2475 ATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAA 2534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1119 ATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAA 1178

Qy   2535 GGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGC 2594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1179 GGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGC 1238

Qy   2595 AGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACA 2654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1239 AGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACA 1298

Qy   2655 GCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCAC 2714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1299 GCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCAC 1358

Qy   2715 TCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAG 2774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1359 TCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAG 1418

Qy   2775 CACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGG 2834
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1419 CACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGG 1478

Qy   2835 CGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCC 2894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1479 CGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCC 1538

Qy   2895 GTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCG 2954
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1539 GTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCG 1598

Qy   2955 CGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCA 3014
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1599 CGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCA 1658

Qy   3015 TCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCAGCTGC 3074
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1659 TCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCAGCTGC 1718

Qy   3075 GGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACC 3134
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1719 GGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACC 1778

Qy   3135 CGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGA 3194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1779 CGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGA 1838

Qy   3195 AGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGT---------GGCGGCGG 3245
          ||||||||||||||||||||||||||||||||||||||||||||         ||||||||
Db    1839 AGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGTGGCGGCGGCGGCGG 1898

Qy   3246 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAGCTGT 3305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1899 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGAAGCTGT 1958

Qy   3306 AGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCAC 3365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1959 AGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCAC 2018

Qy   3366 CGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCAC 3425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2019 CGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCAC 2078

Qy   3426 TTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACAT 3485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2079 TTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACAT 2138

Qy   3486 GCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAA 3545
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2139 GCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAA 2198

Qy   3546 GACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACATGTGG 3605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2199 GACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACATGTGG 2258

Qy   3606 AAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAG 3665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2259 AAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAG 2318

Qy   3666 CAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCG 3725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2319 CAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCG 2378

Qy   3726 GAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCT 3785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2379 GAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCT 2438

Qy   3786 GAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCA 3845
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2439 GAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCA 2498

Qy   3846 GAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATGTCCT 3905
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2499 GAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATGTCCT 2558

Qy   3906 GGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTT 3965
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2559 GGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTT 2618

Qy   3966 TGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAA 4025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2619 TGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAA 2678
```

```
Qy   4026 GTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCAT 4085
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2679 GTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCAT 2738

Qy   4086 TCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGT 4145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2739 TCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGT 2798

Qy   4146 CAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAA 4205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2799 CAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAA 2858

Qy   4206 GTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCT 4265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2859 GTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCT 2918

Qy   4266 CCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCC 4325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2919 CCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCC 2978

Qy   4326 AGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGA 4385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2979 AGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGA 3038

Qy   4386 ACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTA 4445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3039 ACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTA 3098

Qy   4446 CCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCAC 4505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3099 CCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCAC 3158

Qy   4506 TTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGA 4565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3159 TTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGA 3218

Qy   4566 GATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCA 4625
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3219 GATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCA 3278

Qy   4626 CACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATG 4685
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3279 CACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATG 3338

Qy   4686 TCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTAT 4745
          ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db   3339 TCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGTTTAATTTCCTCTAT 3398

Qy   4746 TGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTCTC 4805
          ||||||||||||||||||          ||||||||||||||||||||||||||||||||
Db   3399 TGATGTACAGTCTGTCATG----------GAATTCTATTTGCTGGGCTTTTTTTTTCTC 3447

Qy   4806 TTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGAC 4865
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3448 TTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGAC 3507

Qy   4866 TTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTG 4925
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3508 TTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTG 3567

Qy   4926 TGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCA 4985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3568 TGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCA 3627

Qy   4986 GCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGG 5045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3628 GCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGG 3687

Qy   5046 GGAAATCAAAACAAAAA 5062
          |||||||||||||||||
Db   3688 GGAAATCAAAACAAAAA 3704
```

```
RESULT 4
AAQ12001
ID   AAQ12001 standard; cDNA; 3715 BP.
XX
AC   AAQ12001;
XX
DT   20-AUG-1991  (first entry)
XX
DE   Full-length human androgen receptor coding sequence.
XX
KW   hAR; DNA-binding protein; steroid hormone; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             532..3288
FT                   /*tag= a
FT                   /product= full-length 918 residue hAR
FT                   /note= "includes shorter 734 residue hAR"
XX
PN   WO9107423-A.
XX
PD   30-MAY-1991.
XX
PF   19-OCT-1990;   90WO-US06015.
XX
PR   17-NOV-1989;   89US-0438775.
XX
PA   (ARCH-) ARCH DEV CORP.
XX
PI   Liao S,  Chang C;
XX
DR   WPI; 1991-178048/24.
DR   P-PSDB; AAR12223.
XX
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
XX
PS   Claim 4; Fig 3; 79pp; English.
XX
CC   This sequence was isolated by screening commercially available human
CC   testis and prostate lambda gt11 cDNA libraries. Initial screening
CC   was with probes designed for homology to nucleotide sequences in the
CC   DNA-binding domain of known steroid receptors. Positive clones were
CC   then screened with 24mer probes specific for the various steroid
CC   receptors to eliminate those which coded for known receptors. Any
CC   remaining clones were analysed by restriction mapping and were
CC   sequenced. The human AR coding sequence is given here.
XX
SQ   Sequence 3715 BP; 842 A; 1055 C; 1003 G; 815 T; 0 other;
```

Query Match 70.7%; Score 3594.2; DB 12; Length 3715;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3684; Conservative 0; Mismatches 8; Indels 45; Gaps 4;

```
Qy   1335 TGCACGCGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCC 1394
          | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4 TTCCGGCGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCC 63

Qy   1395 CACCCCGAGTGCGGAGCCAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCA 1454
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     64 CACCCCGAGTGCGGAG-CAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCA 122

Qy   1455 AAAAACAAAACAAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTCAGTT 1514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 AAAAACAAAACAAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTCAGTT 182

Qy   1515 CTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTC 1574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 CTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTC 242

Qy   1575 TTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGC 1634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    243 TTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGC 302

Qy   1635 CTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGGCTGTCA 1694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 CTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGCCTGTCA 362

Qy   1695 GAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGG 1754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 GAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGG 422

Qy   1755 CAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAA 1814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 CAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAA 482

Qy   1815 GGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCA 1874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 GGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCA 542

Qy   1875 GTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCA 1934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 GTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCA 602

Qy   1935 GAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGC 1994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 GAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGC 662

Qy   1995 CGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGCAGCAGCAGCAGCA 2054
          ||||||||||||||||||||||||||||||||||||||||
Db    663 CGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCT---------------------- 701

Qy   2055 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGAC 2114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 ---GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGAC 758

Qy   2115 TAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAG 2174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    759 TAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAG 818

Qy   2175 AGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGC 2234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    819 AGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGC 878

Qy   2235 CCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAG 2294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    879 CCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAG 938

Qy   2295 CAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATC 2354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    939 CAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATC 998

Qy   2355 CACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAA 2414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    999 CACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAA 1058

Qy   2415 AGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGT 2474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1059 AGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGT 1118

Qy   2475 ATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAA 2534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1119 ATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAA 1178

Qy   2535 GGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGC 2594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1179 GGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGC 1238

Qy   2595 AGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACA 2654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1239 AGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACA 1298

Qy   2655 GCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCAC 2714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1299 GCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCAC 1358

Qy   2715 TCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAG 2774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1359 TCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAG 1418

Qy   2775 CACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGCCTAGAAGG 2834
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1419 CACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGG 1478

Qy   2835 CGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCC 2894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1479 CGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCC 1538

Qy   2895 GTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCG 2954
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1539 GTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCG 1598

Qy   2955 CGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCA 3014
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1599 CGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCA 1658

Qy   3015 TCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGC 3074
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1659 TCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGC 1718

Qy   3075 GGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACC 3134
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1719 GGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGCGACC 1778

Qy   3135 CGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGA 3194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1779 CGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGA 1838

Qy   3195 AGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGT---------GGCGGCGG 3245
          ||||||||||||||||||||||||||||||||||||||||||||         ||||||||
Db   1839 AGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGTGGCGGCGGCGGCGG 1898

Qy   3246 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAGCTGT 3305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1899 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAAGCTGT 1958

Qy   3306 AGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCAC 3365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1959 AGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCAC 2018

Qy   3366 CGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCAC 3425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2019 CGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCAC 2078

Qy   3426 TTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACAT 3485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2079 TTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACAT 2138

Qy   3486 GCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAA 3545
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2139 GCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAA 2198

Qy   3546 GACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACATGTGG 3605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2199 GACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACATGTGG 2258

Qy   3606 AAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAG 3665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2259 AAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAG 2318

Qy   3666 CAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCG 3725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2319 CAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCG 2378

Qy   3726 GAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCT 3785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2379 GAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCT 2438
```

```
Qy   3786 GAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCA 3845
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2439 GAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCA 2498

Qy   3846 GAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATGTCCT 3905
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2499 GAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATGTCCT 2558

Qy   3906 GGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTT 3965
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2559 GGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTT 2618

Qy   3966 TGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAA 4025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2619 TGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAA 2678

Qy   4026 GTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCAT 4085
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2679 GTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCAT 2738

Qy   4086 TCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGT 4145
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2739 TGAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGT 2798

Qy   4146 CAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAA 4205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2799 CAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAA 2858

Qy   4206 GTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCT 4265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2859 GTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCT 2918

Qy   4266 CCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCC 4325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2919 CCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCC 2978

Qy   4326 AGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGA 4385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2979 AGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGA 3038

Qy   4386 ACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTA 4445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3039 ACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTA 3098

Qy   4446 CCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCAC 4505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3099 CCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCAC 3158

Qy   4506 TTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGA 4565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3159 TTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGA 3218

Qy   4566 GATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCA 4625
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3219 GATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCA 3278

Qy   4626 CACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATG 4685
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3279 CACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATG 3338

Qy   4686 TCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTAT 4745
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3339 TCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTAT 3398

Qy   4746 TGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTCTC 4805
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   3399 TGATGTACAGTCTGTCATG----------GAATTCTATTTGCTGGGCTTTTTTTTTCTC 3447

Qy   4806 TTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGAC 4865
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3448 TTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGAC 3507

Qy   4866 TTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTG 4925
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3508 TTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTG 3567

Qy   4926 TGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCA 4985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3568 TGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCA 3627

Qy   4986 GCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGG 5045
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   3628 GCCACACAAACGTTTACTTATATTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGG 3687

Qy   5046 GGAAATCAAAACAAAAA 5062
          |||||||||||||||||
Db   3688 GGAAATCAAAACAAAAA 3704
```

RESULT 5
AAT63407
ID AAT63407 standard; cDNA; 3569 BP.
XX
AC AAT63407;
XX
DT 22-JUN-1997 (first entry)
XX
DE Androgen receptor cDNA.
XX
KW Androgen receptor; acidic fibroblast growth factor; aFGF;
KW antisense; benign prostatic hyperplasia; prostate cancer; therapy;
KW ds.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	CDS	363..3122
FT		/*tag= a
FT	misc_feature	complement (916..936)
FT		/*tag= b
FT		/note= "antisense oligonucleotide preferred for
FT		use in methods of the invention"
FT	misc_feature	complement (927..947)
FT		/*tag= c
FT		/note= "antisense oligonucleotide preferred for
FT		use in methods of the invention"
FT	misc_feature	complement (927..936)
FT		/*tag= d
FT		/note= "antisense oligonucleotide preferred for
FT		use in methods of the invention"

XX
PN WO9711170-A1.
XX
PD 27-MAR-1997.
XX
PF 20-SEP-1996; 96WO-US15081.
XX
PR 20-SEP-1995; 95US-0004018.
XX
PA (WORC-) WORCESTER FOUND BIOMEDICAL RES.
XX
PI Zamecnik PA;
XX
DR WPI; 1997-202879/18.
DR P-PSDB; AAW14783.
XX
PT Oligonucleotide(s) antisense to human androgen receptor and acidic
PT FGF genes - used to inhibit gene expression, for the treatment of
PT benign prostatic hyperplasia
XX
PS Claim 2; Page 21-29; 51pp; English.
XX
CC A cDNA clone (AAT63407) codes for the human androgen receptor
CC (AAW14783). Methods of selectively inhibiting the growth, or of
CC killing, prostatic cells involve the use of antisense
CC oligonucleotides (see also AAT63200, AAT63404-05) to this androgen

CC receptor sequence or antisense oligonucleotides (see also AAT63406)
CC to the human acidic fibroblast growth factor gene (see also
CC AAT63197-99). The methods are esp. useful for the treatment of
CC benign prostatic hyperplasia and prostate cancer.
XX
SQ Sequence 3569 BP; 796 A; 1009 C; 974 G; 790 T; 0 other;

Query Match 69.8%; Score 3547; DB 18; Length 3569;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3569; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

```
Qy   1502 TAATAACTCAGTTCTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGA 1561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TAATAACTCAGTTCTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGA 60

Qy   1562 TTTTGTTTTTTTCTTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAG 1621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTTTGTTTTTTTCTTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAG 120

Qy   1622 ACAGACTGTGAGCCTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGAC 1681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACAGACTGTGAGCCTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGAC 180

Qy   1682 TTTGAGGCTGTCAGAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCT 1741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTTGAGGCTGTCAGAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCT 240

Qy   1742 TCCCGCAGGTGGGCAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTT 1801
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCCCGCAGGTGGGCAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTT 300

Qy   1802 CTGAGCAAGAGAAGGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAA 1861
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGAGCAAGAGAAGGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAA 360

Qy   1862 GGATGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACC 1921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGATGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACC 420

Qy   1922 GAGGAGCTTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCA 1981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAGGAGCTTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCA 480

Qy   1982 GGCACCCAGAGGCCGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGC 2041
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGCACCCAGAGGCCGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGC 540

Qy   2042 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2101
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    541 AGCAGCAGCAGCAGCA----------GCAGCAGCAGCAGCAGCAGCAGCAGCAGC 588

Qy   2102 AGCAGCAAGAGACTAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCC 2161
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGCAGCAAGAGACTAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCC 648

Qy   2162 AAGCCCATCGTAGAGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCAC 2221
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 AAGCCCATCGTAGAGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCAC 708

Qy   2222 AGCCGCAGTCGGCCCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCG 2281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 AGCCGCAGTCGGCCCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCG 768

Qy   2282 CCGTGGCCGCCAGCAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACT 2341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CCGTGGCCGCCAGCAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACT 828

Qy   2342 CAGCTGCCCCATCCACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCT 2401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 CAGCTGCCCCATCCACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCT 888

Qy   2402 CCGCTGACCTTAAAGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGC 2461
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    889 CCGCTGACCTTAAAGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGC 948

Qy   2462 AGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTC 2521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    949 AGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTC 1008

Qy   2522 CCACTTCCTCCAAGGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGG 2581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1009 CCACTTCCTCCAAGGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGG 1068

Qy   2582 AGTTGTGTAAGGCAGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGA 2641
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1069 AGTTGTGTAAGGCAGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGA 1128

Qy   2642 GTCCAGGGGAACAGCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCG 2701
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1129 GTCCAGGGGAACAGCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCG 1188

Qy   2702 CTGTGCGTCCCACTCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACA 2761
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1189 CTGTGCGTCCCACTCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACA 1248

Qy   2762 GCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCA 2821
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1249 GCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCA 1308

Qy   2822 AAGGGCTAGAAGGCGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGA 2881
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1309 AAGGGCTAGAAGGCGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGA 1368

Qy   2882 CACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTG 2941
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1369 CACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTG 1428

Qy   2942 CGTACCAGAGTCGCGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTC 3001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1429 CGTACCAGAGTCGCGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTC 1488

Qy   3002 CGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCG 3061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1489 CGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCG 1548

Qy   3062 CCTGGGCGGCTGCGGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGG 3121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1549 CCTGGGCGGCTGCGGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGG 1608

Qy   3122 GTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTC 3181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1609 GTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTC 1668

Qy   3182 TCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGCG 3241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1669 TCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGCG 1728

Qy   3242 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAG 3301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1729 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAG 1788

Qy   3302 CTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACT 3361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1789 CTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACT 1848

Qy   3362 TCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTC 3421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1849 TCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTC 1908

Qy   3422 CCACTTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGG 3481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1909 CCACTTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGG 1968

Qy   3482 ACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCC 3541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1969 ACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCC 2028
```

```
Qy   3542 AGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACAT 3601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2029 AGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACAT 2088

Qy   3602 GTGGAAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCG 3661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2089 GTGGAAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCG 2148

Qy   3662 CCAGCAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTC 3721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2149 CCAGCAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTC 2208

Qy   3722 TTCGGAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTA 3781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2209 TTCGGAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTA 2268

Qy   3782 ATCTGAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAA 3841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2269 ATCTGAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAA 2328

Qy   3842 CCCAGAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATG 3901
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2329 CCCAGAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATG 2388

Qy   3902 TCCTGGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACT 3961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2389 TCCTGGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACT 2448

Qy   3962 CCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGG 4021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2449 CCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGG 2508

Qy   4022 TCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTG 4081
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2509 TCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTG 2568

Qy   4082 TCATTCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCA 4141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2569 TCATTCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCA 2628

Qy   4142 ATGTCAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGC 4201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2629 ATGTCAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGC 2688

Qy   4202 ACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGAT 4261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2689 ACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGAT 2748

Qy   4262 GGCTCCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTA 4321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2749 GGCTCCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTA 2808

Qy   4322 TTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCA 4381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2809 TTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCA 2868

Qy   4382 AGGAACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCT 4441
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2869 AGGAACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCT 2928

Qy   4442 TCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGT 4501
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2929 TCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGT 2988

Qy   4502 TCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGG 4561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2989 TCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGG 3048

Qy   4562 CAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATT 4621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3049 CAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATT 3108

Qy   4622 TCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCA 4681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3109 TCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCA 3168

Qy   4682 GATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCT 4741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3169 GATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCT 3228

Qy   4742 CTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTT 4801
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3229 CTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTT 3288

Qy   4802 TCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTC 4861
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3289 TCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTC 3348

Qy   4862 AGACTTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAA 4921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3349 AGACTTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAA 3408

Qy   4922 TCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGT 4981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3409 TCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGT 3468

Qy   4982 GCCAGCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTAT 5041
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3469 GCCAGCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTAT 3528

Qy   5042 CTGGGGAAATCAAAACAAAAAACAAGCAAACAAAAAAAAAA 5082
          ||||||||||||||||||||||||||||||||||||||||
Db   3529 CTGGGGAAATCAAAACAAAAAACAAGCAAACAAAAAAAAAA 3569
```

RESULT 6
AAN91772
ID AAN91772 standard; cDNA; 3569 BP.
XX
AC AAN91772;
XX
DT 19-MAR-1990 (first entry)
XX
DE Human androgen receptor cDNA.
XX
KW Human androgen receptor; monoclonal antibody; ploYclonal antibody;
KW cancer; probe.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 363..3122
FT /*tag= a
XX
PN WO8909791-A.
XX
PD 19-OCT-1989.
XX
PF 13-APR-1989; 89WO-US01548.
XX
PR 14-APR-1988; 88US-0182646.
XX
PA (UYNC-) UNIVERSITY OF NORTH CAROLINA.
XX
PI French FS, Wilson EM, Joseph DR, Lubahn DB;
XX
DR WPI; 1989-324206/44.
DR P-PSDB; AAP93109.
XX
PT DNA encoding androgen receptor protein - useful for transforming
PT eukaryotic hosts for protein expression and subsequent antibody prodn.
XX
PS Disclosure; Fig. 4; 41pp; English.
XX
CC Complementary DNA sequences derived from the cDNA may be used as probes

```
CC   detect the presence of androgen receptor (AR mRNA in tumour cells, and to
CC   detect AR gene defects using DNA hybrisidation assays.
XX
SQ   Sequence 3569 BP; 796 A; 1008 C; 975 G; 790 T; 0 other;

  Query Match             69.8%;  Score 3545.4;  DB 10;  Length 3569;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 3568;  Conservative    0;  Mismatches    1;  Indels   12;  Gaps    1;

Qy   1502 TAATAACTCAGTTCTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGA 1561
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TAATAACTCAGTTCTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGA 60

Qy   1562 TTTTGTTTTTTTCTTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAG 1621
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTTTGTTTTTTTCTTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAG 120

Qy   1622 ACAGACTGTGAGCCTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGAC 1681
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACAGACTGTGAGCCTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGAC 180

Qy   1682 TTTGAGGCTGTCAGAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCT 1741
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTTGAGGCTGTCAGAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCT 240

Qy   1742 TCCCGCAGGTGGGCAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTT 1801
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCCCGCAGGTGGGCAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTT 300

Qy   1802 CTGAGCAAGAGAAGGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAA 1861
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGAGCAAGAGAAGGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAA 360

Qy   1862 GGATGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACC 1921
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGATGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACC 420

Qy   1922 GAGGAGCTTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCA 1981
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAGGAGCTTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCA 480

Qy   1982 GGCACCCAGAGGCCGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGC 2041
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGCACCCAGAGGCCGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGC 540

Qy   2042 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2101
           ||||||||||||||||||              ||||||||||||||||||||||||||||||
Db    541 AGCAGCAGCAGCAGCA------------GCAGCAGCAGCAGCAGCAGCAGCAGC 588

Qy   2102 AGCAGCAAGAGACTAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCC 2161
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGCAGCAAGAGACTAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCC 648

Qy   2162 AAGCCCATCGTAGAGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCAC 2221
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 AAGCCCATCGTAGAGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCAC 708

Qy   2222 AGCCGCAGTCGGCCCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCG 2281
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 AGCCGCAGTCGGCCCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCG 768

Qy   2282 CCGTGGCCGCCAGCAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACT 2341
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CCGTGGCCGCCAGCAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACT 828

Qy   2342 CAGCTGCCCCATCCACGTTGTCCCTGCTGGGCCCCACTTTCCCCCGGCTTAAGCAGCTGCT 2401
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 CAGCTGCCCCATCCACGTTGTCCCTGCTGGGCCCCACTTTCCCCCGGCTTAAGCAGCTGCT 888

Qy   2402 CCGCTGACCTTAAAGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGC 2461
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 CCGCTGACCTTAAAGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGC 948

Qy   2462 AGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTC 2521
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    949 AGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTC 1008

Qy   2522 CCACTTCCTCCAAGGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGG 2581
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1009 CCACTTCCTCCAAGGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGG 1068

Qy   2582 AGTTGTGTAAGGCAGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGA 2641
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1069 AGTTGTGTAAGGCAGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGA 1128

Qy   2642 GTCCAGGGGAACAGCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCG 2701
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1129 GTCCAGGGGAACAGCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCG 1188

Qy   2702 CTGTGCGTCCCACTCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACA 2761
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1189 CTGTGCGTCCCACTCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACA 1248

Qy   2762 GCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCA 2821
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1249 GCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCA 1308

Qy   2822 AAGGGCTAGAAGGCGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGA 2881
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1309 AAGGGCTAGAAGGCGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGA 1368

Qy   2882 CACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTG 2941
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1369 CACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTG 1428

Qy   2942 CGTACCAGAGTCGCGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTC 3001
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1429 CGTACCAGAGTCGCGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTC 1488

Qy   3002 CGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCG 3061
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1489 CGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCG 1548

Qy   3062 CCTGGGCGGCTGCGGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGG 3121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1549 CCTGGGCGGCTGCGGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGG 1608

Qy   3122 GTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTC 3181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1609 GTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTC 1668

Qy   3182 TCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGCG 3241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1669 TCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGCG 1728

Qy   3242 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAG 3301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1729 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAG 1788

Qy   3302 CTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACT 3361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1789 CTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACT 1848

Qy   3362 TCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTC 3421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1849 TCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTC 1908

Qy   3422 CCACTTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGG 3481
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1909 CCACTTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGG 1968

Qy   3482 ACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCC 3541
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1969 ACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCC 2028

Qy   3542 AGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACAT 3601
```

```
Db  2029 AGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACAT 2088
Qy  3602 GTGGAAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCG 3661
Db  2089 GTGGAAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCG 2148
Qy  3662 CCAGCAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTC 3721
Db  2149 CCAGCAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTC 2208
Qy  3722 TTCGGAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTA 3781
Db  2209 TTCGGAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTA 2268
Qy  3782 ATCTGAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAA 3841
Db  2269 ATCTGAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAA 2328
Qy  3842 CCCAGAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATG 3901
Db  2329 CCCAGAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGGCCATCTTTCTGAATG 2388
Qy  3902 TCCTGGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACT 3961
Db  2389 TCCTGGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACT 2448
Qy  3962 CCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGG 4021
Db  2449 CCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGG 2508
Qy  4022 TCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTG 4081
Db  2509 TCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTG 2568
Qy  4082 TCATTCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCA 4141
Db  2569 TCATTCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCA 2628
Qy  4142 ATGTCAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGC 4201
Db  2629 ATGTCAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGC 2688
Qy  4202 ACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGAT 4261
Db  2689 ACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGAT 2748
Qy  4262 GGCTCCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTA 4321
Db  2749 GGCTCCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTA 2808
Qy  4322 TTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCA 4381
Db  2809 TTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCA 2868
Qy  4382 AGGAACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCT 4441
Db  2869 AGGAACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCT 2928
Qy  4442 TCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGT 4501
Db  2929 TCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGT 2988
Qy  4502 TCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGG 4561
Db  2989 TCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGG 3048
Qy  4562 CAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATT 4621
Db  3049 CAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATT 3108
Qy  4622 TCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCA 4681
Db  3109 TCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCA 3168
Qy  4682 GATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCT 4741
Db  3169 GATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCT 3228
Qy  4742 CTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTT 4801
Db  3229 CTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTT 3288
Qy  4802 TCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTC 4861
Db  3289 TCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTC 3348
Qy  4862 AGACTTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAA 4921
Db  3349 AGACTTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAA 3408
Qy  4922 TCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGT 4981
Db  3409 TCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGT 3468
Qy  4982 GCCAGCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTAT 5041
Db  3469 GCCAGCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTAT 3528
Qy  5042 CTGGGGAAATCAAAACAAAAAACAAGCAAACAAAAAAAAAA 5082
Db  3529 CTGGGGAAATCAAAACAAAAAACAAGCAAACAAAAAAAAAA 3569
```

RESULT 7
AAS93472/c
ID AAS93472 standard; cDNA; 3590 BP.
XX
AC AAS93472;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #29276.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG29285.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 29276; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The

```
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3590 BP; 807 A; 977 C; 1016 G; 790 T; 0 other;

  Query Match             68.8%;  Score 3498.2;  DB 23;  Length 3590;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 3565;  Conservative    0;  Mismatches    3;  Indels   25;  Gaps    4;

Qy   1502 TAATAACTCAGTTCTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGA 1561
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3590 TAATAACTCAGTTCTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGA 3531

Qy   1562 TTTTGTTTTTTTCTTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAG 1621
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3530 TTTTGTTTTTTTCTTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAG 3471

Qy   1622 ACAGACTGTGAGCCTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGAC 1681
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3470 ACAGACTGTGAGCCTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGAC 3411

Qy   1682 TTTGAGGCTGTCAGAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCT 1741
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3410 TTTGAGGCTGTCAGAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCT 3351

Qy   1742 TCCCGCAGGTGGGCAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTT 1801
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3350 TCCCGCAGGTGGGCAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTT 3291

Qy   1802 CTGAGCAAGAGAAGGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAA 1861
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3290 CTGAGCAAGAGAAGGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAA 3231

Qy   1862 GGATGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACC 1921
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3230 GGATGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACC 3171

Qy   1922 GAGGAGCTTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCA 1981
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3170 GAGGAGCTTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCA 3111

Qy   1982 GGCACCCAGAGGCCGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGC 2041
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   3110 GGCACCCAGAGGCCGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCA-CAGC 3052

Qy   2042 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2101
           |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3051 AGC--CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2994

Qy   2102 AGCAGCAAGAGACTAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCC 2161
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2993 AGCAGCAAGAGACTAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCC 2934

Qy   2162 AAGCCCATCGTAGAGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCAC 2221
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2933 AAGCCCATCGTAGAGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCAC 2874

Qy   2222 AGCCGCAGTCGGCCCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCG 2281
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2873 AGCCGCAGTCGGCCCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCG 2814

Qy   2282 CCGTGGCCGCCAGCAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACCAGGATGACT 2341
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2813 CCGTGGCCGCCAGCAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACCAGGATGACT 2754

Qy   2342 CAGCTGCCCCATCCACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCT 2401
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2753 CAGCTGCCCCATCCACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCT 2694

Qy   2402 CCGCTGACCTTAAAGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGC 2461
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2693 CCGCTGACCTTAAAGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGC 2634

Qy   2462 AGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTC 2521
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2633 AGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTC 2574

Qy   2522 CCACTTCCTCCAAGGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGG 2581
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2573 CCACTTCCTCCAAGGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGG 2514

Qy   2582 AGTTGTGTAAGGCAGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGA 2641
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2513 AGTTGTGTAAGGCAGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGA 2454

Qy   2642 GTCCAGGGGAACAGCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCG 2701
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2453 GTCCAGGGGAACAGCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCG 2394

Qy   2702 CTGTGCGTCCCACTCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACA 2761
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2393 CTGTGCGTCCCACTCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACA 2334

Qy   2762 GCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCA 2821
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2333 GCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCA 2274

Qy   2822 AAGGGCTAGAAGGCGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGA 2881
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2273 AAGGGCTAGAAGGCGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGA 2214

Qy   2882 CACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTG 2941
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2213 CACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTG 2154

Qy   2942 CGTACCAGAGTCGCGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTC 3001
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2153 CGTACCAGAGTCGCGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTC 2094

Qy   3002 CGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCG 3061
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2093 CGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCG 2034

Qy   3062 CCTGGGCGGCTGCGGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGG 3121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2033 CCTGGGCGGCTGCGGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGG 1974

Qy   3122 GTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTC 3181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1973 GTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTC 1914

Qy   3182 TCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGCG 3241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1913 TCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGCG 1854

Qy   3242 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAG 3301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1853 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAG 1794

Qy   3302 CTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACT 3361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

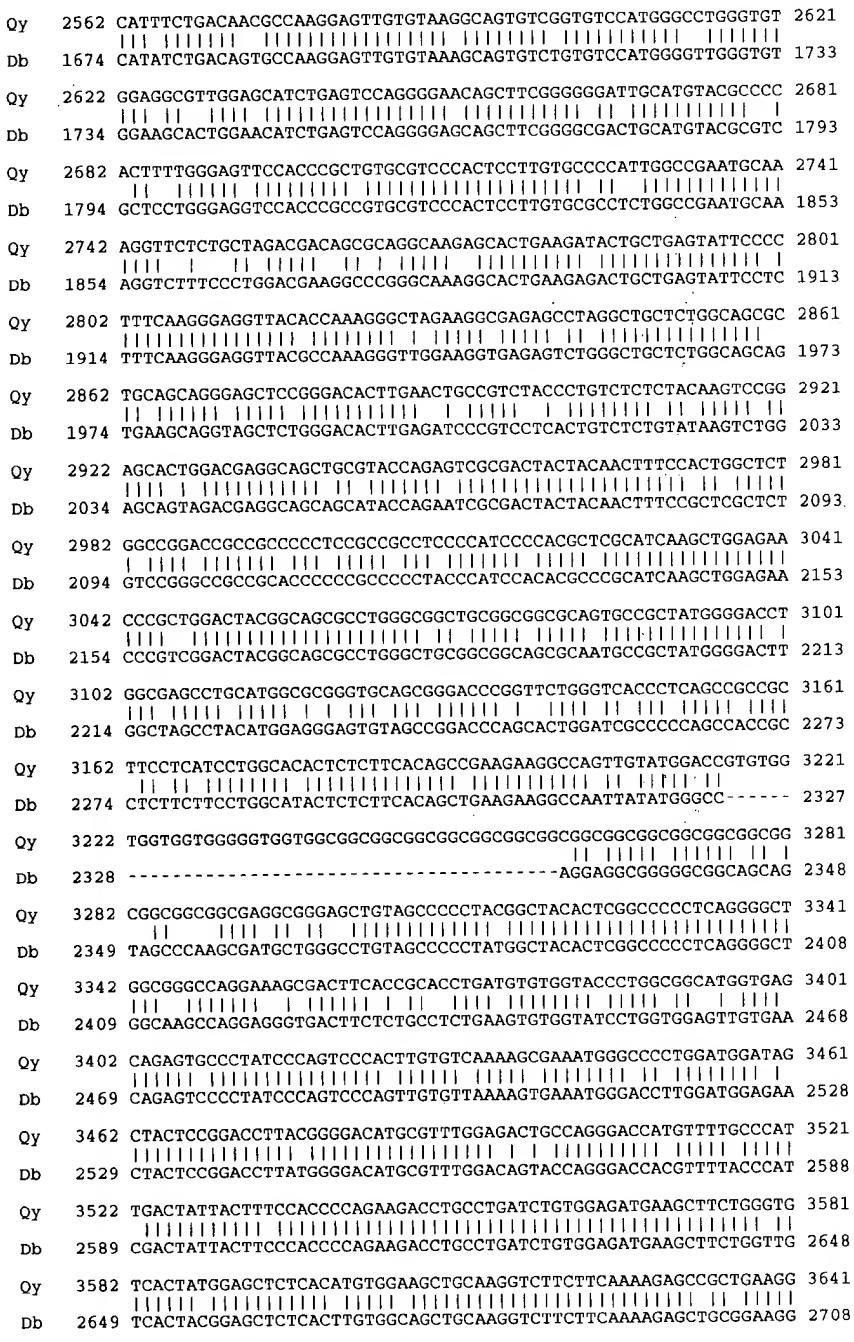
```
Db   1793 CTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACT 1734

Qy   3362 TCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTC 3421

Db   1733 TCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTC 1674

Qy   3422 CCACTTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGG 3481

Db   1673 CCACTTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGG 1614

Qy   3482 ACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCC 3541

Db   1613 ACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCC 1554

Qy   3542 AGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACAT 3601

Db   1553 AGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACAT 1494

Qy   3602 GTGGAAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCG 3661

Db   1493 GTGGAAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCG 1434

Qy   3662 CCAGCAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTC 3721

Db   1433 CCAGCAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTC 1374

Qy   3722 TTCGGAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTA 3781

Db   1373 TTCGGAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTA 1314

Qy   3782 ATCTGAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAA 3841

Db   1313 ATCTGAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAA 1254

Qy   3842 CCCAGAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATG 3901

Db   1253 CCCAGAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATG 1194

Qy   3902 TCCTGGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACT 3961

Db   1193 TCCTGGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACT 1134

Qy   3962 CCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGG 4021

Db   1133 CCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGG 1074

Qy   4022 TCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTG 4081

Db   1073 TCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTG 1014

Qy   4082 TCATTCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCA 4141

Db   1013 TCATTCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCA 954

Qy   4142 ATGTCAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGC 4201

Db    953 ATGTCAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGC 894

Qy   4202 ACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGAT 4261

Db    893 ACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGAT 834

Qy   4262 GGCTCCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTA---------- 4308

Db    833 GGCTCCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCCATGCTAATTTTTTTTTTT 774

Qy   4309 --------CTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATG 4360

Db    773 TTTTGCTTCTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATG 714

Qy   4361 AACTTCGAATGAACTACATCAAGGAACTCGATCGTATCATTGCATGCAAAAGAAAAAATC 4420

Db    713 AACTTCGAATGAACTACATCAAGGAACTCGATCGTATCATTGCATGCAAAAGAAAAAATC 654

Qy   4421 CCACATCCTGCTCAAGACGCTTCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTA 4480

Db    653 CCACATCCTGCTCAAGACGCTTCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTA 594

Qy   4481 TTGCGAGAGAGCTGCATCAGTTCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCG 4540

Db    593 TTGCGAGAGAGCTGCATCAGTTCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCG 534

Qy   4541 TGGACTTTCCGGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTG 4600

Db    533 TGGACTTTCCGGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTG 474

Qy   4601 GGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCAC 4660

Db    473 GGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCAC 414

Qy   4661 CCCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTG 4720

Db    413 CCCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTG 354

Qy   4721 CAGTGCCTTGGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATT 4780

Db    353 CAGTGCCTTGGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATT 294

Qy   4781 CTATTTGCTGGGCTTTTTTTTTCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATC 4840

Db    293 CTATCTGCTGGGCTTTTTTTTTCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATC 234

Qy   4841 TAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGA 4900

Db    233 TAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGA 174

Qy   4901 ATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGTGC 4960

Db    173 ATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGTGC 114

Qy   4961 TTGTTTACAGCACTACTCTGTGCCAGCCACACAAA-CGTTTACTTATCTTATGCCACGGG 5019

Db    113 TTGTTTACAGCACTACTCTGTGCCAGCCACACAAACCGTTTACTTATCTTATGCCACGGG 54

Qy   5020 AAGTTTAGAGAGCTAAGATTATCTGGGGAAATCAAAACAAAAAACAAGCAAAC 5072

Db     53 AAGTTTAGAGAGCTAAGATTATCTGGGGAAATCAAAACAAAAAACAAGCAAAC 1
```

RESULT 8
AAN91773
ID AAN91773 standard; cDNA; 4180 BP.
XX
AC AAN91773;
XX
DT 19-MAR-1990 (first entry)
XX
DE Rat androgen receptor cDNA.
XX
KW Rat androgen receptor; monoclonal antibody; ploYclonal antibody;
KW cancer; probe.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
FT CDS 996..3702
FT /*tag= a
XX
PN WO8909791-A.
XX
PD 19-OCT-1989.
XX
PF 13-APR-1989; 89WO-US01548.
XX
PR 14-APR-1988; 88US-0182646.
XX

PA (UYNC-) UNIVERSITY OF NORTH CAROLINA.
XX
PI French FS, Wilson EM, Joseph DR, Lubahn DB;
XX
DR WPI; 1989-324206/44.
DR P-PSDB; AAP93110.
XX
PT DNA encoding androgen receptor protein - useful for transforming
PT eukaryotic hosts for protein expression and subsequent antibody prodn.
XX
PS Disclosure; Fig. 5; 41pp; English.
XX
CC Complementary DNA sequences derived from the cDNA may be used as probes
CC detect the presence of androgen receptor (AR) mRNA in tumour cells, and
CC to detect AR gene defects using DNA hybrisidation assays.
XX
SQ Sequence 4180 BP; 1024 A; 1149 C; 1083 G; 924 T; 0 other;

Query Match 48.9%; Score 2486.4; DB 10; Length 4180;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 3416; Conservative 0; Mismatches 664; Indels 306; Gaps 23;

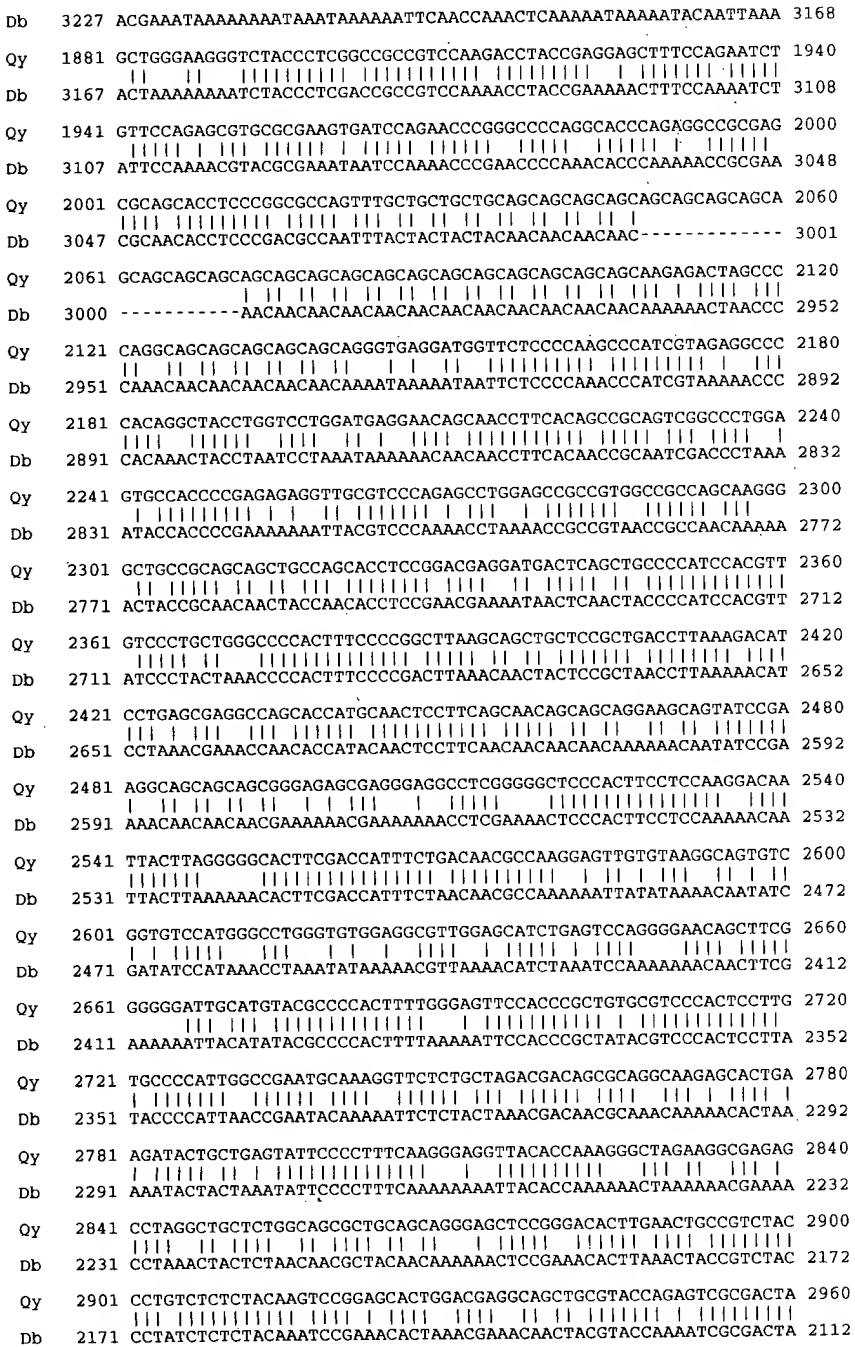
```
Qy   758 ATCCCGGGGAGCCAGCTTGCTGGGAGAGCGGGAACGGTCCGGAGCAAGCCCAGAGGCAGA 817
Db     1 ATCCCTAGGAGCCAGCCTGCTGGGAGAACCAG-AGGGTCCGGAGCAAACCTGGAGGCTGA 59

Qy   818 GGAGGCGACAGAGGGAAAAAGGGCCCNAGCTAGCCGCTCCAGTGCTGTACAGNAGCCGAA 877
Db    60 GAGGGCATCAGAGGGGAAAAGA--CTGAGTTAGCCACTCCAGTGCCATACAGAAGCTTAA 117

Qy   878 -GGACGCACCACGCCAGCCCCAGCCCGGCTCCAGCGACAGCNAACGCCTCTTGCANGCGT 936
Db   118 GGGACATACCACGCCAGCCCCAGCCCAGCGACAGCCAACGCCTGTTGCAGAGCGGCGGCT 177

Qy   937 TCGAAGCCGCCGCCC-GGAGCTGCCCTTTCCTCTTCGGTGAAGTTTTTAAAAGCTGCTAA 995
Db   178 TCGAAGCCGCCGCCCAGAAGCTGCCCTTTCCTCTTCGGTGAAGTTTCTAAAAGCTGCGGG 237

Qy   996 AGACTCGGAGGAAGCAAGGAAAGTGCCTGGTAGGACTGACGGCTGCCTTTGTCCTCCTCC 1055
Db   238 AGACTCGGAGGAAGCGAAGAAAGTGTCCGGTAGGACT-ACGACTGCCTTTGTCCTCCTCC 296

Qy  1056 TCTCCACCCCGCCTCCCCCCACCCTGCCTTCCCCCCCTCCCCCGTCTTCTCTCCCGCAGC 1115
Db   297 CTCCTACCC----------CTACCCCTCCTGGGTCCCCTCTCCCTGAGC 335

Qy  1116 TGCCTCAGTCGGCTACTCTCAGCCAACCCCCCTCACCACCCTTCTCCCCACCCGCCCCCC 1175
Db   336 GGACTAGGCAGGCTTC----------CTGGCCAGCC 361

Qy  1176 CGCCCCCGTCGGCCCAGCGNTGNCAGNCCGAGTTTGCAGAGAGGTAACTCCCTTTGGCTG 1235
Db   362 CTCTCCCCTACACCACCAGCTCTGCCAGCCAGTTTGCACAGAGGTAACTCCCTTTGGCTG 421

Qy  1236 CGAGCGGGCGAGNCTAGCTGCACATTGCAAAGAAGGCTCTTAGGAG-CAGGCGACTGGGG 1294
Db   422 AAAGCAGACGAG-CTTGTTGCCCATTGGAAGGGAGGCTTTTGGGAGCCCAGAGACTGAGG 480

Qy  1295 AGCGGCTTCAGCACTGCAGCCACGACCNGCCTGGTTAGGCTGCACGCGGAGAGAACCCTC 1354
Db   481 AGCAACAGCACGCTGGAGAGTCCCTGATTCCAGGTT---------- 516

Qy  1355 TGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCC--CACCCCGAGTGCGGAGCC 1412
Db   517 ----------CTCCCCCCTGCACCTCCTACTGCCCGCCCCTCACCCTGTGTGTGCAGCT 565

Qy  1413 AGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCAAAAAACAAAACAAACAAA 1472
Db   566 AGAATTGAAAAGATGAAAAGACAGTTGGGGCTTCAGTAGTCGAAAGCAAAAC-------- 617

Qy  1473 AACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTCAGTTCTTATTTGCACCTACTTC 1532
Db   618 -----AAAAGCAAAAAGAAAACAAAAAGAAAATAGCCCAGTTCTTATTTGCACCTGCTTC 672

Qy  1533 AGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTCTTTTAAGATCTGGGCATC 1592
Db   673 AGTGGACATTGACTTTGGAAGGCAGAGAATTTTCCTTCCCCCCAGTCAAGCTTTGAGCAT 732

Qy  1593 TTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGCCTAGCAGGGCAGATCTTG 1652
Db   733 CTTTTAATCTGTTCTTCAAGTATTTAGGGACAAACTGTGAAACTAGCAGGGCAGATCCTG 792

Qy  1653 TCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGGCTGTCAGAGCGCTTTTTGCGTGGT 1712
Db   793 TCTAGCGCGTGCCTTCCTTTACAGGAGACTTTGAGGCTATCTGGGCGCT-----CCCCCC 847

Qy  1713 TGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGGCAGCTAGCTGCAGCGACT 1772
Db   848 CCTCCCTGCAAGTTTTCTTCCCTGGAGCTTCCCGCAGGTGGGCAGCTAGCTGCAGATACT 907

Qy  1773 ACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAAGGGGAGGCGGGGTAAGGG 1832
Db   908 ACATCAT-----CAGTCAGTAGAACTCTTCAGAGCAAGAGACGAGGAGGCAGGATAAGGG 962

Qy  1833 AAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCAGTTAGGGCTGGGAAGGGT 1892
Db   963 AATTCGGTGGAAGCTAGAGACAAGCTAAAGGATGGAGGTGCAGTTAGGGCTGGGAAGGGT 1022

Qy  1893 CTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCAGAATCTGTTCCAGAGCGT 1952
Db   1023 CTACCCACGGCCCCCGTCCAAGACCTATCGAGGAGCGTTCCAGAATCTGTTCCAGAGCGT 1082

Qy  1953 GCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGCCGCGAGCGCAGCACCTCC 2012
Db   1083 GCGCGAAGCGATCCAGAACCCGGGCCCCAGGCACCCTGAGGCCGCTAGCATAGCACCTCC 1142

Qy  2013 CGGCGCCAGTTTGCTGCTGCTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 2072
Db   1143 CGGTGCCTGTTTACAGCAGCGGCAGGAGACTAGCCCCCGGC---------- 1183

Qy  2073 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGACTAGCCCCAGGCAGCAGCA 2132
Db   1184 ----------GGCGGCGGCG 1193

Qy  2133 GCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAGAGGCCCCACAGGCTACCT 2192
Db   1194 GCAGCAGCACCCTGAGGATGGCTCTCCTCAAGCCCACATCAGAGGCACCACAGGCTACCT 1253

Qy  2193 GGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGCCCTGGAGTGCCACCCCGA 2252
Db   1254 GGCCCTGGAGGAGGAACAGCAGCCTTCACAGCAGCAGTCAGCCTCCGAGGGCCACCCTGA 1313

Qy  2253 GAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAGCAAGGGGCTGCCGCAGCA 2312
Db   1314 GAGCGGCTGCCTCCCGGAGCCTGGAGCTGCCACGGCTCCTGGCAAGGGGCTGCCGCAGCA 1373

Qy  2313 GCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATCCACGTTGTCCCTGCTGGG 2372
Db   1374 GCCACCAGCTCCTCCAGATCAGGATGACTCAGCTGCCCCATCCACGTTGTCCCTACTGGG 1433

Qy  2373 CCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAAAGACATCCTGAGCGAGGC 2432
Db   1434 CCCCACTTTCCCAGGCTTAAGCAGCTGCTCCGCAGACATTAAAGACATCCTGAGCGAGGC 1493

Qy  2433 CAGCACCATGCAACTCCTT---------- 2451
Db   1494 CGGCACCATGCAACTTCTTCAGCAGCAGCAGCAACAGCAACAGCAGCAGCAGCAGCAGCA 1553

Qy  2452 ----------CAGCAACAGCAGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGAGAGC 2501
Db   1554 GCAGCAGCAGCAGCAACAGCAGCAGGAGGTAATATCCGAAGGCAGCAGCAGCGTGAGAGC 1613

Qy  2502 GAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAAGGACAATTACTTAGGGGGCACTTCGAC 2561
Db   1614 AAGGGAGGCCACTGGGGCTCCCTCTTCCTCCAAGGATAGTTACCTAGGGGGCAATTCGAC 1673
```


Qy 2562 CATTTCTGACAACGCCAAGGAGTTGTGTAAGGCAGTGTCGGTGTCCATGGGCCTGGGTGT 2621
Db 1674 CATATCTGACAGTGCCAAGGAGTTGTGTAAAGCAGTGTCTGTGTCCATGGGGTTGGGTGT 1733
Qy 2622 GGAGGCGTTGGAGCATCTGAGTCCAGGGGAACAGCTTCGGGGGGATTGCATGTACGCCCC 2681
Db 1734 GGAAGCACTGGAACATCTGAGTCCAGGGGAGCAGCTTCGGGGCGACTGCATGTACGCGTC 1793
Qy 2682 ACTTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCCTTGTGCCCCATTGGCCGAATGCAA 2741
Db 1794 GCTCCTGGGAGGTCCACCCGCCGTGCGTCCCACTCCTTGTGCGCCTCTGGCCGAATGCAA 1853
Qy 2742 AGGTTCTCTGCTAGACGACAGCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCC 2801
Db 1854 AGGTCTTTCCCTGGACGAAGGCCCGGGCAAAGGCACTGAAGAGACTGCTGAGTATTCCTC 1913
Qy 2802 TTTCAAGGGAGGTTACACCAAAGGGCTAGAAGGCGAGAGCCTAGGCTGCTCTGGCAGCGC 2861
Db 1914 TTTCAAGGGAGGTTACGCCAAAGGGTTGGAAGGTGAGAGTCTGGGCTGCTCTGGCAGCAG 1973
Qy 2862 TGCAGCAGGGAGCTCCGGGACACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGG 2921
Db 1974 TGAAGCAGGTAGCTCTGGGACACTTGAGATCCCGTCCTCACTGTCTCTGTATAAGTCTGG 2033
Qy 2922 AGCACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACTACTACAACTTTCCACTGGCTCT 2981
Db 2034 AGCAGTAGACGAGGCAGCAGCATACCAGAATCGCGACTACTACAACTTTCCGCTCGCTCT 2093
Qy 2982 GGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAA 3041
Db 2094 GTCCGGGCCGCCGCACCCCCCGCCCCCTACCCATCCACACGCCCGCATCAAGCTGGAGAA 2153
Qy 3042 CCCGCTGGACTACGGCAGCGCCTGGGCGGCTGCGGCGGCGCAGTGCCGCTATGGGGACCT 3101
Db 2154 CCCGTCGGACTACGGCAGCGCCTGGGCTGCGGCGGCAGCGCAATGCCGCTATGGGGACTT 2213
Qy 3102 GGCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCGC 3161
Db 2214 GGCTAGCCTACATGGAGGGAGTGTAGCCGGACCCAGCACTGGATCGCCCCCAGCCACCGC 2273
Qy 3162 TTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGTGTGG 3221
Db 2274 CTCTTCTTCCTGGCATACTCTCTTCACAGCTGAAGAAGGCCAATTATATGGGCC------ 2327
Qy 3222 TGGTGGTGGGGGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3281
Db 2328 ----------AGGAGGCGGGGGCGGCAGCAG 2348
Qy 3282 CGGCGGCGGCGAGGCGGGAGCTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCT 3341
Db 2349 TAGCCCAAGCGATGCTGGGCCTGTAGCCCCCTATGGCTACACTCGGCCCCCTCAGGGGCT 2408
Qy 3342 GGCGGGCCAGGAAAGCGACTTCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAG 3401
Db 2409 GGCAAGCCAGGAGGGTGACTTCTCTGCCTCTGAAGTGTGGTATCCTGGTGGAGTTGTGAA 2468
Qy 3402 CAGAGTGCCCTATCCCAGTCCCACTTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAG 3461
Db 2469 CAGAGTCCCCTATCCCAGTCCCAGTTGTGTTAAAAGTGAAATGGGACCTTGGATGGAGAA 2528
Qy 3462 CTACTCCGGACCTTACGGGGACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCAT 3521
Db 2529 CTACTCCGGACCTTATGGGGACATGCGTTTGGACAGTACCAGGGACCACGTTTTACCCAT 2588
Qy 3522 TGACTATTACTTTCCACCCCAGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTG 3581
Db 2589 CGACTATTACTTCCCACCCCAGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGTTG 2648
Qy 3582 TCACTATGGAGCTCTCACATGTGGAAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGG 3641
Db 2649 TCACTACGGAGCTCTCACTTGTGGCAGCTGCAAGGTCTTCTTCAAAAGAGCTGCGGAAGG 2708
Qy 3642 GAAACAGAAGTACCTGTGCGCCAGCAGAAATGATTGCACTATTGATAAATTCCGAAGGAA 3701
Db 2709 GAAACAGAAGTATCTATGTGCCAGCAGAAATGATTGCACCATTGATAAATTTCGGAGGAA 2768
Qy 3702 AAATTGTCCATCTTGTCGTCTTCGGAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCG 3761
Db 2769 AAATTGTCCATCGTGTCGTCTCCGGAAATGTTATGAAGCAGGGATGACTCTGGGAGCTCG 2828
Qy 3762 GAAGCTGAAGAAACTTGGTAATCTGAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCAC 3821
Db 2829 TAAGCTGAAGAAACTTGGAAATCTCAAACTACAGGAAGAAGGAGAAAACTCCAGTGCTGG 2888
Qy 3822 CAGCCCCACTGAGGAGACAACCCAGAAGCTGACAGTGTCACACATTGAAGGCTATGAATG 3881
Db 2889 TAGCCCCACTGAGGACCCATCCCAGAAGATGACTGTATCACACATTGAAGGCTATGAATG 2948
Qy 3882 TCAGCCCATCTTTCTGAATGTCCTGGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACA 3941
Db 2949 TCAACCTATCTTTCTTAATGTCCTGGAAGCCATTGAGCCAGGAGTGGTGTGTGCCGGACA 3008
Qy 3942 CGACAACAACCAGCCCGACTCCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGA 4001
Db 3009 TGACAACAACCAGCCTGATTCCTTTGCTGCCTTGTTATCTAGTCTCAACGAGCTTGGCGA 3068
Qy 4002 GAGACAGCTTGTACACGTGGTCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACA 4061
Db 3069 GAGACAGCTTGTACATGTGGTCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTGCA 3128
Qy 4062 CGTGGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCAT 4121
Db 3129 TGTGGATGACCAGATGGCAGTCATTCAGTATTCCTGGATGGGACTGATGGTATTTGCCAT 3188
Qy 4122 GGGCTGGCGATCCTTCACCAATGTCAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGT 4181
Db 3189 GGGTTGGCGGTCCTTCACTAATGTCAACTCTAGGATGCTCTACTTTGCACCTGACCTGGT 3248
Qy 4182 TTTCAATGAGTACCGCATGCACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCA 4241
Db 3249 TTTCAATGAGTATCGCATGCACAAGTCTCGAATGTACAGCCAGTGCGTGAGGATGAGGCA 3308
Qy 4242 CCTCTCTCAAGAGTTTGGATGGCTCCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGC 4301
Db 3309 CCTTTCTCAAGAGTTTGGATGGCTCCAGATAACCCCCCAGGAATTCCTGTGCATGAAAGC 3368
Qy 4302 ACTGCTACTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGA 4361
Db 3369 ACTGCTACTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGA 3428
Qy 4362 ACTTCGAATGAACTACATCAAGGAACTCGATCGTATCATTGCATGCAAAAGAAAAAATCC 4421
Db 3429 ACTTCGAATGAACTACATCAAGGAACTTGATCGCATCATTGCATGCAAAAGAAAAAATCC 3488
Qy 4422 CACATCCTGCTCAAGACGCTTCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTAT 4481
Db 3489 CACATCCTGCTCAAGGCGCTTCTACCAGCTCACCAAGCTCCTGGATTCTGTGCAGCCTAT 3548
Qy 4482 TGCGAGAGAGCTGCATCAGTTCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCGT 4541
Db 3549 TGCAAGAGAGCTGCATCAATTCACTTTTGACCTGCTAATCAAGTCCCATATGGTGAGCGT 3608
Qy 4542 GGACTTTCCGGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGG 4601
Db 3609 GGACTTTCCTGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGG 3668
Qy 4602 GAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCACC 4661
Db 3669 GAAAGTCAAGCCCATCTATTTCCACACACAGTGAAGATTTGGAAACCCTAATACCCAAAC 3728
Qy 4662 CCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTGT--TATAACTCTGCACTACTCCTCT 4719
Db 3729 CCACCTTGT-TCCCTTTTCAGATGTCTTCTGCCTGTTATATAACTCTGCACTACTCTCT 3787
Qy 4720 GCAGTGCCTT-GGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAACATGTTCCTGAA 4778

```
Db   3788 GCATTGCCTTGGGGGAAATTCCTCTACTGATGTACAGTCTGTCATGAACATGTTCCCCAA 3847

Qy   4779 TTCTATTTGCTGGGCTTTTTTTTTCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTA 4838
                 |  || |||  |   ||||  || ||||||||||||||||| || |||
Db   3848 ----------GTTCTATTTCCTGGGCTTTTCCTTCTTTCTTTTTCTTCTTCTCTGCCTC 3896

Qy   4839 TCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCTATCTGTGTTTT 4898
          | | ||||||||||||||| ||  || |  ||| |   |||||||||| |||||||||
Db   3897 TTTTACCCTCCCATGGCACATTTTGAATCCGCTGCGTGTTGTGGCTCCTGCCTGTGTTTT 3956

Qy   4899 GAATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGT 4958
          || |  |||||||  ||| || |||||||||||| || | |||||||||||||  |||
Db   3957 GAGTTTTGTTGTATTTCTTCAAGTCTGTGATGATCTTCTTGTGGCCCAGTGTCAA-CTGT 4015

Qy   4959 GCTTGTTTACAGCAC--TACTCTGTGCCAGCCACACAAACGTTTACTTATCTTATGCCAC 5016
          |||||||||| |||||  | || ||||||| |||  |||| ||||||| | |||||||||
Db   4016 GCTTGTTTATAGCACTGTGCTGTGTGCCAACCAAGCAAATGTTTACTCACCTTATGCCAT 4075

Qy   5017 GGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAATCAAAACAAAAAACAAGCAAACAAAA 5076
          || ||||||||||||||  | ||||| |||||   | ||  | | |   ||| ||
Db   4076 GGCAAGTTTAGAGAGCTATAAGTATCTTGGGAAGAAACAAACAGAGAGAGTAAAAAAACC 4135

Qy   5077 AAAAAA 5082
          ||||||
Db   4136 AAAAAA 4141
```

```
RESULT 9
ABK34013/c
ID   ABK34013 standard; DNA; 3715 BP.
XX
AC   ABK34013;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human DNA for staging of Astrocytomas, complement, #50.
XX
KW   Human; ds; astrocytoma; cytostatic; staging; cysteine methylation; CpG;
KW   bisulphite; brain tissue; MALDI; ESI; electron spray mass spectrometry;
KW   matrix assisted laser desorption/ionization mass spectrometry.
XX
OS   Homo sapiens.
XX
PN   WO200202808-A2.
XX
PD   10-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP07538.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-171649/22.
XX
PT   Novel chemically modified genomic DNA sequences, useful in the
PT   characterisation, classification, differentiation, grading, staging,
PT   treatment and/or diagnosis of astrocytomas or predisposition to
PT   astrocytomas -
XX
PS   Claim 1; SEQ ID No 100; 37pp; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence (I) of at
CC   least 18 bases in length of a segment of chemically pre-treated genomic
CC   DNA which has any one of the sequences of (ABK33919-ABK34032) or its
CC   complement. Also included are an oligonucleotide or peptide nucleic
CC   acid (or set thereof) of at least 9 nucleotides which hybridises to (I),
CC   primers for (I), probes for detecting cytosine methylation or single-
CC   nucleotide polymorphisms (SNP) in (I), an array of oligomers
CC   or peptide nucleic acids for analysing diseases associated with the
CC   methylation states of the CpG dinucleotides of (I). The array is useful
CC   for determining genetic and/or epigenetic parameters, classification,
CC   differentiation, grading, staging, treatment and/or diagnosis of
CC   astrocytomas, or the predisposition to astrocytomas by analysing cytosine
CC   methylations, involves obtaining a biological sample containing genomic
CC   DNA, extracting the genomic DNA, converting cytosine bases which are
CC   unmethylated at the 5-position, in the genomic DNA sample, to uracil or
CC   another base which is dissimilar to cytosine in terms of hybridisation
CC   behaviour, by chemical treatment and amplifying chemically pre-treated
CC   genomic DNA fragments using the array and a polymerase, where the
CC   amplificates carry a detectable label. The method further involves
CC   identifying methylation status of one or more cytosine positions, and
CC   analysing methylation status of the cytosine positions by reference to
CC   one or more data sets. The genomic DNA is chemically treated by using a
CC   bisulphite, hydrogen sulphite or disulphite. The amplification
CC   step amplifies DNA which is of particular interest in astrocytoma or
CC   brain tissue, based on the specific genomic methylation status of brain
CC   tissues, as opposed to background DNA. The amplificates carry a
CC   fluorescent label or radionuclide. Optionally, the labels of the
CC   amplificates are detachable molecule fragments having a typical mass
CC   which are detected in a mass spectrometer. The fragments of chemically
CC   pre-treated genomic DNA to be amplified, have a single positive or
CC   negative charge for a better detectability in the mass spectrometer.
CC   Preferably, the amplificates or fragments of the amplificates are
CC   detected by matrix assisted laser desorption/ionization mass spectrometry
CC   (MALDI) or using electron spray mass spectrometry (ESI). The
CC   present sequence is one of the chemically pre-treated reference DNA
CC   samples of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3715 BP; 818 A; 149 C; 1055 G; 1693 T; 0 other;
```

```
  Query Match           44.2%;  Score 2244.2;  DB 24;  Length 3715;
  Best Local Similarity 76.1%;  Pred. No. 0;
  Matches 2838;  Conservative    0;  Mismatches  848;  Indels   45;  Gaps    4;

Qy   1341 CGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCCCACCCC 1400
          || | | |||||||| ||||||||||||||||||||||||||||||||| |||||||||||
Db   3706 CGAAAAAAACCCTCTATTTTCCCCCACTCTCTCTCCACCTCCTCCTACCTTCCCCCACCCC 3647

Qy   1401 GAGTGCGGAGCCAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCAAAAAAC 1460
          || | | ||  || | ||||||| || ||||  || |||  |||||| || |||||||||
Db   3646 GAATAC-GAAACAAAAATCAAAAAATAAAAAAACAATCAAATCTTCAATAACCAAAAAAC 3588

Qy   1461 AAAACAAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTCAGTTCTTATT 1520
          |||||||||||||||||||| |||||||||| ||||| |||||||||| |||||||||
Db   3587 AAAACAAACAAAAACAAAAAAACCGAAATAAAAAAAAAAAATAATAACTCAATTCTTATT 3528

Qy   1521 TGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTCTTTTAA 1580
          | |||||||||||  ||||| ||||  ||  |  |  ||||| ||||||||||||||
Db   3527 TACACCTACTTCAATAAACACTAAATTTAAAAAATAAAAAATTTTATTTTTTTCTTTTAA 3468

Qy   1581 GATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGCCTAGCA 1640
            ||||   |||||||| |||||||||||||| |||||  | ||| ||| | | |||| ||
Db   3467 AATCTAAACATCTTTTAAATCTACCCTTCAAATATTAAAAAACAAACTATAAACCTAACA 3408

Qy   1641 GGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGGCTGTCAGAGCGC 1700
            || ||||| |||||||| | ||||||||| ||||| ||||| |  || ||| | |||
Db   3407 AAACAAATCTTATCCACCGTATATCTTCTTCTACACGAAACTTTAAAACTATCAAAACGC 3348

Qy   1701 TTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGGCAGCTA 1760
          ||||| |||  || |||||||||| |||||||||| | |||||||||  |   || |||
Db   3347 TTTTTACGTAATTACTCCCGCAAATTTCCTTCTCTAAAACTTCCCGCAAATAAACAACTA 3288

Qy   1761 GCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAACAGAAGGGGAG 1820
            || || |||||||||||||||||| ||| || ||||||||| | ||| | ||    |
Db   3287 ACTACAACGACTACCGCATCATCACAACCTATTAAACTCTTCTAAACAAAAAAAAAAAAA 3228

Qy   1821 GCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCAGTTAGG 1880
            ||   |||   || ||  |  || ||||| |||| |||||  ||  || | || |||
```


Db 3227 ACGAAATAAAAAAAATAAATAAAAAATTCAACCAAACTCAAAAATAAAAATACAATTAAA 3168
Qy 1881 GCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCAGAATCT 1940
Db 3167 ACTAAAAAAAATCTACCCTCGACCGCCGTCCAAAACCTACCGAAAAACTTTCCAAAATCT 3108
Qy 1941 GTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGCCGCGAG 2000
Db 3107 ATTCCAAAACGTACGCGAAATAATCCAAAACCCGAACCCAAACACCCAAAAACCGCGAA 3048
Qy 2001 CGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGCAGCAGCAGCAGCAGCA 2060
Db 3047 CGCAACACCTCCCGACGCCAATTTACTACTACTACAACAACAACAAC---------- 3001
Qy 2061 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGACTAGCCC 2120
Db 3000 ----------AACAACAACAACAACAACAACAACAACAACAACAACAAAAAACTAACCC 2952
Qy 2121 CAGGCAGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAGAGGCCC 2180
Db 2951 CAAACAACAACAACAACAACAAAATAAAAATAATTCTCCCCAAACCCATCGTAAAAACCC 2892
Qy 2181 CACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGCCCTGGA 2240
Db 2891 CACAAACTACCTAATCCTAAATAAAAAACAACAACCTTCACAACCGCAATCGACCCTAAA 2832
Qy 2241 GTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAGCAAGGG 2300
Db 2831 ATACCACCCCGAAAAAAATTACGTCCCAAAACCTAAAACCGCCGTAACCGCCAACAAAAA 2772
Qy 2301 GCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATCCACGTT 2360
Db 2771 ACTACCGCAACAACTACCAACACCTCCGAACGAAAATAACTCAACTACCCCATCCACGTT 2712
Qy 2361 GTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAAAGACAT 2420
Db 2711 ATCCCTACTAAACCCCACTTTCCCCGACTTAAACAACTACTCCGCTAACCTTAAAAACAT 2652
Qy 2421 CCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGTATCCGA 2480
Db 2651 CCTAAACGAAACCAACACCATACAACTCCTTCAACAACAACAACAAAAAACAATATCCGA 2592
Qy 2481 AGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAAGGACAA 2540
Db 2591 AAACAACAACAACGAAAAAACGAAAAAAACCTCGAAAACTCCCACTTCCTCCAAAAACAA 2532
Qy 2541 TTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGCAGTGTC 2600
Db 2531 TTACTTAAAAAACACTTCGACCATTTCTAACAACGCCAAAAAATTATATAAAACAATATC 2472
Qy 2601 GGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACAGCTTCG 2660
Db 2471 GATATCCATAAACCTAAATATAAAAACGTTAAAACATCTAAATCCAAAAAAACAACTTCG 2412
Qy 2661 GGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCCTTG 2720
Db 2411 AAAAAATTACATATACGCCCCACTTTTAAAAATTCCACCCGCTATACGTCCCACTCCTTA 2352
Qy 2721 TGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAGCACTGA 2780
Db 2351 TACCCCATTAACCGAATACAAAAATTCTCTACTAAACGACAACGCAAACAAAAACACTAA 2292
Qy 2781 AGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGGCGAGAG 2840
Db 2291 AAATACTACTAAATATTCCCCTTTCAAAAAAAATTACACCAAAAAACTAAAAAACGAAAA 2232
Qy 2841 CCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCCGTCTAC 2900
Db 2231 CCTAAACTACTCTAACAACGCTACAACAAAAAACTCCGAAACACTTAAACTACCGTCTAC 2172
Qy 2901 CCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACTA 2960
Db 2171 CCTATCTCTCTACAAATCCGAAACACTAAACGAAACAACTACGTACCAAAATCGCGACTA 2112
Qy 2961 CTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCCA 3020
Db 2111 CTACAACTTTCCACTAACTCTAACCGAACCGCCGCCCCCTCCGCCGCCTCCCCATCCCCA 2052
Qy 3021 CGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGCGGCGGC 3080
Db 2051 CGCTCGCATCAAACTAAAAAACCCGCTAAACTACGACAACGCCTAAACGACTACGACGAC 1992
Qy 3081 GCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTTC 3140
Db 1991 GCAATACCGCTATAAAAACCTAACGAACCTACATAACGCGAATACAACGAAACCCGATTC 1932
Qy 3141 TGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAGG 3200
Db 1931 TAAATCACCCTCAACCGCCGCTTCCTCATCCTAACACACTCTCTTCACAACCGAAAAAAA 1872
Qy 3201 CCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGT---------GGTGGCGGCGGCGGCGG 3251
Db 1871 CCAATTATATAAACCGTATAATAATAATAAAAATAATAATAACGACGACGACGACGACGA 1812
Qy 3252 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAGCTGTAGCCCC 3311
Db 1811 CGACGACGACGACGACGACGACGACGACGACGACGACGAAACGAAAACTATAACCCC 1752
Qy 3312 CTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCACCGCACC 3371
Db 1751 CTACGACTACACTCGACCCCCTCAAAAACTAACGAACCAAAAAAACGACTTCACCGCACC 1692
Qy 3372 TGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCACTTGTGT 3431
Db 1691 TAATATATAATACCCTAACGACATAATAAACAAAATACCCTATCCCAATCCCACTTATAT 1632
Qy 3432 CAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACATGCGTTT 3491
Db 1631 CAAAAACGAAATAAACCCCTAAATAAATAACTACTCCGAACCTTACGAAAACATACGTTT 1572
Qy 3492 GGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAAGACCTG 3551
Db 1571 AAAAACTACCAAAAACCATATTTTACCCATTAACTATTACTTTCCACCCCAAAAAACCTA 1512
Qy 3552 CCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACATGTGGAAGCTG 3611
Db 1511 CCTAATCTATAAAAATAAAACTTCTAAATATCACTATAAAACTCTCACATATAAAAACTA 1452
Qy 3612 CAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAGCAGAAA 3671
Db 1451 CAAAATCTTCTTCAAAAAAACCGCTAAAAAAAAACAAAAATACCTATACGCCAACAAAAA 1392
Qy 3672 TGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCGGAAATG 3731
Db 1391 TAATTACACTATTAATAAATTCCGAAAAAAAAATTATCCATCTTATCGTCTTCGAAAATA 1332
Qy 3732 TTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCTGAAACT 3791
Db 1331 TTATAAAACAAAAATAACTCTAAAAACCCGAAAACTAAAAAAACTTAATAATCTAAAACT 1272
Qy 3792 ACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCAGAAGCT 3851
Db 1271 ACAAAAAAAAAAAAAAAACTTCCAACACCACCAACCCCACTAAAAAAACAACCCAAAAACT 1212
Qy 3852 GACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATGTCCTGGAAGC 3911
Db 1211 AACAATATCACACATTAAAAACTATAAATATCAACCCATCTTTCTAAATATCCTAAAAAC 1152
Qy 3912 CATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTTTGCAGC 3971
Db 1151 CATTAAACCAAATATAATATATACTAAACACGACAACAACCAACCCGACTCCTTTACAAC 1092
Qy 3972 CTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAAGTGGGC 4031
Db 1091 CTTACTCTCTAACCTCAATAAACTAAAAAAAAAACAACTTATACACGTAATCAAATAAAC 1032

```
Qy   4032 CAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCATTCAGTA 4091
          ||| |||| ||| |||||||||||||||||| ||||||| || || ||||||| ||
Db   1031 CAAAACCTTACCTAACTTCCGCAACTTACACGTAAACGACCAAATAACTATCATTCAATA 972

Qy   4092 CTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGTCAACTC 4151
          ||||| ||    ||||| | ||| |||| || ||||||||||||||| |||||||
Db    971 CTCCTAAATAAAACTCATAATATTTACCATAAACTAACGATCCTTCACCAATATCAACTC 912

Qy   4152 CAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCCG 4211
          ||  || |||||||||||||| |||| |||||||| | |||||||| ||||| |||||
Db    911 CAAAATACTCTACTTCGCCCCTAATCTAATTTTCAATAAATACCGCATACACAAATCCCG 852

Qy   4212 GATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCAAAT 4271
          || |||| ||| | | ||||||| | |||||||||||| | ||| || ||||||||
Db    851 AATATACAACCAATATATCCGAATAAAACACCTCTCTCAAAAATTTAAATAACTCCAAAT 792

Qy   4272 CACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCCAGTGGA 4331
          |||||||||  ||||||| | ||| ||| |||| |||||||||| |||||||||| |  |
Db    791 CACCCCCCAAAAATTCCTATACATAAAAACACTACTACTCTTCAACATTATTCCAATAAA 732

Qy   4332 TGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTCGA 4391
          |  || |||||||||||||||||| || |||||||||| |||||||||| |||||||
Db    731 TAAACTAAAAAATCAAAAATTCTTTAATAAACTTCGAATAAACTACATCAAAAAACTCGA 672

Qy   4392 TCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTACCAGCT 4451
          |||||||||| ||| ||||| |||||||||||||||||| ||||| ||||||||||| ||
Db    671 TCGTATCATTACATACAAAAAAAAAAATCCCACATCCTACTCAAAACGCTTCTACCAACT 612

Qy   4452 CACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCACTTTTGA 4511
          |||||| |||||  ||||||| || |||||| ||| | | || ||||| |||||||||  |
Db    611 CACCAAACTCCTAAACTCCGTACAACCTATTACGAAAAAACTACATCAATTCACTTTTAA 552

Qy   4512 CCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGAGATCAT 4571
          ||| |||||||| |||||||  | | |||  |||||||| |||| ||  || | ||||
Db    551 CCTACTAATCAAATCACACATAATAAACGTAAACTTTCCGAAAATAATAACAAAAATCAT 492

Qy   4572 CTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACCCA 4631
          |||| | ||| | ||||| |||||||||   ||| |||| ||||||||||||||||||||
Db    491 CTCTATACAAATACCCAAAATCCTTTCTAAAAAAATCAAACCCATCTATTTCCACACCCA 432

Qy   4632 GTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATGTCTTCT 4691
          | || ||||  |||||||||||||||||||||| ||||| |||||||||| || ||||||
Db    431 ATAAAACATTAAAAACCCTATTTCCCCACCCCAACTCATACCCCCTTTCAAATATCTTCT 372

Qy   4692 GCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTATTGATGT 4751
          ||| |||||||||| |||||||||| || | ||||    |||||||||||| || |
Db    371 ACCTATTATAACTCTACACTACTCCTCTACAATACCTTATTTAATTTCCTCTATTAATAT 312

Qy   4752 ACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTCTCTTTCTC 4811
          ||| ||| |||| |           |||||||||| ||   ||||||||||||||||||
Db    311 ACAATCTATCATAA----------AATTCTATTTACTAAACTTTTTTTTTCTCTTTCTC 263

Qy   4812 TCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTTGCT 4871
          ||||||||||||||||||||||||||||||||||||||||||||  |||||||| ||||| ||
Db    262 TCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATAACACCTTCAAACTTTACT 203

Qy   4872 TCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATGA 4931
          |||||||  |  ||||||||| | |||| |||  | || ||| |||||||||| | || |
Db    202 TCCCATTATAACTCCTATCTATATTTTAAATAATATTATATACCTTTAAATCTATAATAA 143

Qy   4932 TCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCAGCCACA 4991
          |||||||||  |||| | |||| || | ||| |||||| |||||||||  | ||| ||||
Db    142 TCCTCATATAACCCAATATCAAATTATACTTATTTACAACACTACTCTATACCAACCACA 83

Qy   4992 CAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAAT 5051
          |||||||||||||||||||| |||||  || |||| | | |||| |||||||    ||||
Db     82 CAAACGTTTACTTATCTTATACCACGAAAAATTTAAAAAACTAAAATTATCTAAAAAAAT 23

Qy   5052 CAAAACAAAAA 5062
          ||||||||||
Db     22 CAAAACAAAAA 12
```

RESULT 10
ABK31545/c
ID ABK31545 standard; DNA; 3715 BP.
XX
AC ABK31545;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified complementary DNA #194.
XX
KW Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200926-A2.
XX
PD 03-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP07472.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-147896/19.
XX
PT Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction
XX
PS Claim 1; SEQ ID No 388; 24pp; English.
XX
CC The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or
CC disulphite. Also disclosed are oligonucleotides and/or PNA oligomers
CC for detecting the cytosine methylation state (CpG islands) of these
CC genes, and a method for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with signal transduction.
CC The genomic DNA can be obtained from cells or cellular components which
CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
CC histologic object slides, and all their possible combinations. The
CC sequences of the invention are useful for the diagnosis and therapy of
CC diseases associated with signal transduction e.g. solid tumours and
CC cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA
CC sequences of different genes associated with signal transduction, or
CC their complementary sequences.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 3715 BP; 818 A; 149 C; 1055 G; 1693 T; 0 other;

Query Match 44.2%; Score 2244.2; DB 24; Length 3715;
Best Local Similarity 76.1%; Pred. No. 0;
Matches 2838; Conservative 0; Mismatches 848; Indels 45; Gaps 4;

```
Qy   1341 CGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCCCACCCC 1400
          || | | |||||||| ||||||||||||||||||||||||||||||||| ||||||||||||
Db   3706 CGAAAAAAACCCTCTATTTTCCCCCACTCTCTCTCCACCTCCTCCTACCTTCCCCACCCC 3647

Qy   1401 GAGTGCGGAGCCAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCAAAAAAC 1460
          || | | ||  || | |||||||| || ||||  || |||  |||||| || ||||||||
Db   3646 GAATAC-GAAACAAAAATCAAAAAATAAAAAAACAATCAAATCTTCAATAACCAAAAAAC 3588
```

```
Db   1451 CAAAATCTTCTTCAAAAAAACCGCTAAAAAAAAACAAAAATACCTATACGCCAACAAAAA 1392

Qy   3672 TGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCGGAAATG 3731
Db   1391 TAATTACACTATTAATAAATTCCGAAAAAAAAATTATCCATCTTATCGTCTTCGAAAATA 1332

Qy   3732 TTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCTGAAACT 3791
Db   1331 TTATAAAACAAAAATAACTCTAAAAACCCGAAAACTAAAAAAACTTAATAATCTAAAACT 1272

Qy   3792 ACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCAGAAGCT 3851
Db   1271 ACAAAAAAAAAAAAAAACTTCCAACACCACCAACCCCACTAAAAAAACAACCCAAAAACT 1212

Qy   3852 GACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATGTCCTGGAAGC 3911
Db   1211 AACAATATCACACATTAAAAACTATAAATATCAACCCATCTTTCTAAATATCCTAAAAAC 1152

Qy   3912 CATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTTTGCAGC 3971
Db   1151 CATTAAACCAAATATAATATATACTAAACACGACAACAACCAACCCGACTCCTTTACAAC 1092

Qy   3972 CTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAAGTGGGC 4031
Db   1091 CTTACTCTCTAACCTCAATAAACTAAAAAAAAAACAACTTATACACGTAATCAAATAAAC 1032

Qy   4032 CAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCATTCAGTA 4091
Db   1031 CAAAACCTTACCTAACTTCCGCAACTTACACGTAAACGACCAAATAACTATCATTCAATA 972

Qy   4092 CTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGTCAACTC 4151
Db    971 CTCCTAAATAAAACTCATAATATTTACCATAAACTAACGATCCTTCACCAATATCAACTC 912

Qy   4152 CAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCCG 4211
Db    911 CAAAATACTCTACTTCGCCCCTAATCTAATTTTCAATAAATACCGCATACACAAATCCCG 852

Qy   4212 GATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCAAAT 4271
Db    851 AATATACAACCAATATATCCGAATAAAACACCTCTCTCAAAAATTTAAATAACTCCAAAT 792

Qy   4272 CACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCCAGTGGA 4331
Db    791 CACCCCCCAAAAATTCCTATACATAAAAACACTACTACTCTTCAACATTATTCCAATAAA 732

Qy   4332 TGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTCGA 4391
Db    731 TAAACTAAAAAATCAAAAATTCTTTAATAAACTTCGAATAAACTACATCAAAAAACTCGA 672

Qy   4392 TCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTACCAGCT 4451
Db    671 TCGTATCATTACATACAAAAAAAAAAATCCCACATCCTACTCAAAACGCTTCTACCAACT 612

Qy   4452 CACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCACTTTTGA 4511
Db    611 CACCAAACTCCTAAACTCCGTACAACCTATTACGAAAAAACTACATCAATTCACTTTTAA 552

Qy   4512 CCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGAGATCAT 4571
Db    551 CCTACTAATCAAATCACACATAATAAACGTAAACTTTCCGAAAATAATAACAAAAATCAT 492

Qy   4572 CTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACCCA 4631
Db    491 CTCTATACAAATACCCAAAATCCTTTCTAAAAAAATCAAACCCATCTATTTCCACACCCA 432

Qy   4632 GTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATGTCTTCT 4691
Db    431 ATAAAACATTAAAAACCCTATTTCCCCACCCCAACTCATACCCCCTTTCAAATATCTTCT 372

Qy   4692 GCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTATTGATGT 4751
Db    371 ACCTATTATAACTCTACACTACTCCTCTACAATACCTTATTTAATTTCCTCTATTAATAT 312
```

```
Qy   4752 ACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTCTCTTTCTC 4811
Db    311 ACAATCTATCATAA----------AATTCTATTTACTAAACTTTTTTTTTCTCTTTCTC 263

Qy   4812 TCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTTGCT 4871
Db    262 TCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATAACACCTTCAAACTTTACT 203

Qy   4872 TCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATGA 4931
Db    202 TCCCATTATAACTCCTATCTATATTTTAAATAATATTATATACCTTTAAATCTATAATAA 143

Qy   4932 TCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCAGCCACA 4991
Db    142 TCCTCATATAACCCAATATCAAATTATACTTATTTACAACACTACTCTATACCAACCACA 83

Qy   4992 CAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAAT 5051
Db     82 CAAACGTTTACTTATCTTATACCACGAAAAATTTAAAAAACTAAAATTATCTAAAAAAAT 23

Qy   5052 CAAAACAAAAA 5062
Db     22 CAAAACAAAAA 12
```

```
RESULT 11
AAQ12002
ID   AAQ12002 standard; DNA; 3217 BP.
XX
AC   AAQ12002;
XX
DT   20-AUG-1991  (first entry)
XX
DE   Full-length rat androgen receptor coding sequence.
XX
KW   rAR; DNA-binding protein; steroid hormone; ss.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   CDS             33..2741
FT                   /*tag= a
FT                   /product= full-length (902 amino acids) rAR
FT                   /note= "includes 733 residue shorter rAR protein"
XX
PN   WO9107423-A.
XX
PD   30-MAY-1991.
XX
PF   19-OCT-1990;   90WO-US06015.
XX
PR   17-NOV-1989;   89US-0438775.
XX
PA   (ARCH-) ARCH DEV CORP.
XX
PI   Liao S,  Chang C;
XX
DR   WPI; 1991-178048/24.
DR   P-PSDB; AAR12224.
XX
PT   Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT   and antibodies for detection and quantification methods
XX
PS   Claim 4; Fig 3; 79pp; English.
XX
CC   This sequence was isolated by screening a rat ventral prostate
CC   lambda gt11 library in E.coli Y1090. Initial screening
CC   was with probes designed for homology to nucleotide sequences in the
CC   DNA-binding domain of known steroid receptors. Positive clones were
CC   then screened with 24mer probes specific for the various steroid
CC   receptors to eliminate those which coded for known receptors. Any
CC   remaining clones were analysed by restriction mapping and were
```

```
CC   sequenced. The rat AR coding sequence is given here.
XX
SQ   Sequence 3217 BP; 776 A; 873 C; 843 G; 725 T; 0 other;

  Query Match             43.1%;  Score 2189.2;  DB 12;  Length 3217;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches 2726;  Conservative    0;  Mismatches  393;  Indels  191;  Gaps   11;

Qy   1832 GAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCAGTTAGGGCTGGGAAGGG 1891
Db      1 GAATTCGGTGGAAGCTAGAGACAAGCTAAAGGATGGAGGTGCAGTTAGGGCTGGGAAGGG 60

Qy   1892 TCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCAGAATCTGTTCCAGAGCG 1951
Db     61 TCTACCCACGGCCCCCGTCCAAGACCTATCGAGGAGCGTTCCAGAATCTGTTCCAGAGCG 120

Qy   1952 TGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGCCGCGAGCGCAGCACCTC 2011
Db    121 TGCGCGAAGCGATCCAGAACCCGGGCCCCAGGCACCCTGAGGCCGCTAGCATAGCACCTC 180

Qy   2012 CCGGCGCCAGTTTGCTGCTGCTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2071
Db    181 CCGGTGCCTGTTT--------------------------------------------- 193

Qy   2072 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGACTAGCCCCA---GGCAGC 2128
Db    194 ------------------------ACAGCAGCGGCAGGAGACTAGCCCCAGGCGGCGGC 228

Qy   2129 AGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAGAGGCCCCACAGGCT 2188
Db    229 GGCGGCAGCAGCACCCTGAGGATGGCTCTCCTCAAGCCCACATCAGAGGCACCACAGGCT 288

Qy   2189 ACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGCCCTGGAGTGCCACC 2248
Db    289 ACCTGGCCCTGGAGGAGGAACAGCAGCCTTCACAGCAGCAGTCAGCCTCCGAGGGCCACC 348

Qy   2249 CCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAGCAAGGGGCTGCCGC 2308
Db    349 CTGAGAGCGGCTGCCTCCCGGAGCCTGGAGCTGCCACGGCTCCTGGCAAGGGGCTGCCGC 408

Qy   2309 AGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATCCACGTTGTCCCTGC 2368
Db    409 AGCAGCCACCAGCTCCTCCAGATCAGGATGACTCAGCTGCCCCATCCACGTTGTCCCTAC 468

Qy   2369 TGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAAAGACATCCTGAGCG 2428
Db    469 TGGGCCCCACTTTCCCAGGCTTAAGCAGCTGCTCCGCAGACATTAAAGACATCCTGAGCG 528

Qy   2429 AGGCCAGCACCATGCAACTCCTT----------------------------------- 2451
Db    529 AGGCCGGCACCATGCAACTTCTTCAGCAGCAGCAGCAACAGCAACAGCAGCAGCAGCAGC 588

Qy   2452 -------------CAGCAACAGCAGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGA 2497
Db    589 AGCAGCAGCAGCAGCAGCAACAGCAGCAGGAGGTAATATCCGAAGGCAGCAGCAGCGTGA 648

Qy   2498 GAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAAGGACAATTACTTAGGGGGCACTT 2557
Db    649 GAGCAAGGGAGGCCACTGGGGCTCCCTCTTCCTCCAAGGATAGTTACCTAGGGGGCAATT 708

Qy   2558 CGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGCAGTGTCGGTGTCCATGGGCCTGG 2617
Db    709 CGACCATATCTGACAGTGCCAAGGAGTTGTGTAAAGCAGTGTCTGTGTCCATGGGGTTGG 768

Qy   2618 GTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACAGCTTCGGGGGGATTGCATGTACG 2677
Db    769 GTGTGGAAGCACTGGAACATCTGAGTCCAGGGGAGCAGCTTCGGGGCGACTGCATGTACG 828

Qy   2678 CCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCCTTGTGCCCCATTGGCCGAAT 2737
Db    829 CGTCGCTCCTGGGAGGTCCACCCGCCGTGCGTCCCACTCCTTGTGCGCCTCTGGCCGAAT 888

Qy   2738 GCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATT 2797
Db    889 GCAAAGGTCTTTCCCTGGACGAAGGCCCGGGCAAAGGCACTGAAGAGACTGCTGAGTATT 948

Qy   2798 CCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGGCGAGAGCCTAGGCTGCTCTGGCA 2857
Db    949 CCTCTTTCAAGGGAGGTTACGCCAAAGGGTTGGAAGGTGAGAGTCTGGGCTGCTCTGGCA 1008

Qy   2858 GCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGT 2917
Db   1009 GCAGTGAAGCAGGTAGCTCTGGGACACTTGAGATCCCGTCCTCACTGTCTCTGTATAAGT 1068

Qy   2918 CCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACTACTACAACTTTCCACTGG 2977
Db   1069 CTGGAGCAGTAGACGAGGCAGCAGCATACCAGAATCGCGACTACTACAACTTTCCGCTCG 1128

Qy   2978 CTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGG 3037
Db   1129 CTCTGTCCGGGCCGCCGCACCCCCCGCCCCCTACCCATCCACACGCCCGCATCAAGCTGG 1188

Qy   3038 AGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGCGGCGGCGCAGTGCCGCTATGGGG 3097
Db   1189 AGAACCCGTTGGACTACGGCAGCGCCTGGGCTGCGGCGGCAGCGCAATGCCGCTATGGGG 1248

Qy   3098 ACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCG 3157
Db   1249 ACTTGGCTAGCCTACATGGAGGGAGTGTAGCCCGACCCAGCACTGGATCGCCCCCAGCCA 1308

Qy   3158 CCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGT 3217
Db   1309 CCGCCTCTTCTTCCTGGCATACTCTCTTCACAGCTGAAGAAGGCCAGTTGTATGGGCC-- 1366

Qy   3218 GTGGTGGTGGTGGGGGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3277
Db   1367 ------------------------------------------AGGAGGCGGGGGCGGCA 1383

Qy   3278 GCGGCGGCGGCGGCGAGGCGGGAGCTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGG 3337
Db   1384 GCAGTAGCCCAAGCGATGCTGGGCCTGTAGCCCCCTATGGCTACACTCGGCCCCCTCAGG 1443

Qy   3338 GGCTGGCGGGCCAGGAAAGCGACTTCACCGCACCTGATGTGTGGTACCCTGGCGGCATGG 3397
Db   1444 GGCTGGCAAGCCAGGAGGGTGACTTCTCTGCCTCTGAAGTGTGGTATCCTGGTGGAGTTG 1503

Qy   3398 TGAGCAGAGTGCCCTATCCCAGTCCCACTTGTGTCAAAAGCGAAATGGGCCCCTGGATGG 3457
Db   1504 TGAACAGAGTCCCCTATCCCAGTCCCAGTTGTGTTAAAAGTGAAATGGGACCTTGGATGG 1563

Qy   3458 ATAGCTACTCCGGACCTTACGGGGACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGC 3517
Db   1564 AGAACTACTCCGGACCTTATGGGGACATGCGTTTGGACAGTACCAGGGACCACGTTTTAC 1623

Qy   3518 CCATTGACTATTACTTTCCACCCCAGAAGACCTGCCTGATCTGTGGAGATCAAGCTTCTG 3577
Db   1624 CCATCGACTATTACTTCCCACCCCAGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTG 1683

Qy   3578 GGTGTCACTATGGAGCTCTCACATGTGGAAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTG 3637
Db   1684 GTTGTCACTACGGAGCTCTCACTTGTGGCAGCTGCAAGGTCTTCTTCAAAAGAGCTGCGG 1743

Qy   3638 AAGGGAAACAGAAGTACCTGTGCGCCAGCAGAAATGATTGCACTATTGATAAATTCCGAA 3697
Db   1744 AAGGGAAACAGAAGTATCTATGTGCCAGCAGAAATGATTGCACCATTGATAAATTTCGGA 1803

Qy   3698 GGAAAAATTGTCCATCTTGTCGTCTTCGGAAATGTTATGAAGCAGGGATGACTCTGGGAG 3757
Db   1804 GGAAAAATTGTCCATCGTGTCGTCTCCGGAAATGTTATGAAGCAGGGATGACTCTGGGAG 1863

Qy   3758 CCCGGAAGCTGAAGAAACTTGGTAATCTGAAACTACAGGAGGAAGGAGAGGCTTCCAGCA 3817
Db   1864 CCCGGAAGCTGAAGAAACTTGGAAATCTCAAACTACAGGAAGAAGGAGAAAACTCCAGTG 1923

Qy   3818 CCACCAGCCCCACTGAGGAGACAACCCAGAAGCTGACAGTGTCACACATTGAAGGCTATG 3877
```

```
Db   1924 CTGGTAGCCCCACTGAGGACCCATCCCAGAAGATGACTGTATCACACATTGAAGGCTATG 1983

Qy   3878 AATGTCAGCCCATCTTTCTGAATGTCCTGGAAGCCATTGAGCCAGGTGTAGTGTGTGCTG 3937

Db   1984 AATGTCAACCTATCTTTCTTAATGTCCTGGAAGCCATTGAGCCAGGAGTGGTGTGTGCCG 2043

Qy   3938 GACACGACAACAACCAGCCCGACTCCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACTGG 3997

Db   2044 GACATGACAACAACCAGCCTGATTCCTTTGCTGCCTTGTTATCTAGTCTCAACGAGCTTG 2103

Qy   3998 GAGAGAGACAGCTTGTACACGTGGTCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACT 4057

Db   2104 GAGAGAGACAGCTTGTACATGTGGTCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACT 2163

Qy   4058 TACACGTGGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGGCTCATGGTGTTTG 4117

Db   2164 TGCATGTGGATGACCAGATGGCAGTCATTGAGTATTCCTGGATGGGACTGATGGTATTTG 2223

Qy   4118 CCATGGGCTGGCGATCCTTCACCAATGTCAACTCCAGGATGCTCTACTTCGCCCCTGATC 4177

Db   2224 CCATGGGTTGGCGGTCCTTCACTAATGTCAACTCTAGGATGCTCTACTTTGCACCTGACC 2283

Qy   4178 TGGTTTTCAATGAGTACCGCATGCACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGA 4237

Db   2284 TGGTTTTCAATGAGTATCGCATGCACAAGTCTCGAATGTACAGCCAGTGCGTGAGGATGA 2343

Qy   4238 GGCACCTCTCTCAAGAGTTTGGATGGCTCCAAATCACCCCCCAGGAATTCCTGTGCATGA 4297

Db   2344 GGCACCTTTCTCAAGAGTTTGGATGGCTCCAGATAACCCCCCAGGAATTCCTGTGCATGA 2403

Qy   4298 AAGCACTGCTACTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTG 4357

Db   2404 AAGCACTGCTACTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTG 2463

Qy   4358 ATGAACTTCGAATGAACTACATCAAGGAACTCGATCGTATCATTGCATGCAAAAGAAAAA 4417

Db   2464 ATGAACTTCGAATGAACTACATCAAGGAACTTGATCGCATCATTGCATGCAAAAGAAAAA 2523

Qy   4418 ATCCCACATCCTGCTCAAGACGCTTCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGC 4477

Db   2524 ATCCCACATCCTGCTCAAGGCGCTTCTACCAGCTCACCAAGCTCCTGGATTCTGTGCAGC 2583

Qy   4478 CTATTGCGAGAGAGCTGCATCAGTTCACTTTTGACCTGCTAATCAAGTCACACATGGTGA 4537

Db   2584 CTATTGCAAGAGAGCTGCAACAGTTCACTTTTGACCTGCTAATCAAGTCCCATATGGTGA 2643

Qy   4538 GCGTGGACTTTCCGGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTT 4597

Db   2644 GCGTGGACTTTCCTGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTT 2703

Qy   4598 CTGGGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAAACCCTATTTCCC 4657

Db   2704 CTGGGAAAGTCAAGCCCATCTATTTCCACACACAGTGAAGATTTGGAAACCCTAATACCC 2763

Qy   4658 CACCCCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTGT--TATAACTCTGCACTACTC 4715

Db   2764 AAACCCACCTTGT-TCCCTTTTCAGATGTCTTCTGCCTGTTATATAACTCTGCACTACTT 2822

Qy   4716 CTCTGCAGTGCCTT-GGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAACATGTTCC 4774

Db   2823 CTCTGCAGTGCCTTGGGGGAAATTCCTCTACTGATGTACAGTCTGTCATGAACATGTTCC 2882

Qy   4775 TGAATTCTATTTGCTGGGCTTTTTTTTTCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTC 4834

Db   2883 CCAGTTCTATTTCCTGGGCTTTT----------CCTTCTTTCTTTTTCTTCTTCTCTG 2930

Qy   4835 CCTATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCTATCTGTG 4894

Db   2931 CCTCTTTTACCCTCCCATGGCACATTTTGAATCCGCTGCGTGTTGT-GCTCCTGCCTGTG 2989

Qy   4895 TTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCCAGTGTCAAG 4954

Db   2990 TTTTGAGTTTTGTTGTATTTCTTCAAGTCTGTGATGATCTTCTTGTGGCCCAGTGTCAA- 3048

Qy   4955 TTGTGCTTGTTTACAGCAC--TACTCTGTGCCAGCCACACAAACGTTTACTTATCTTATG 5012

Db   3049 CTGTGCTTGTTTATAGCACTGTGCTGTGTGCCAACCAAGCAAATGTTTACTCACCTTATG 3108

Qy   5013 CCACGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAATCAAAACAAAAAACAAGCAAAC 5072

Db   3109 CCATGGCAAGTTTAGAGAGCTATAAGTATCTTGGGAAGAAACAAACAGAGAGAGTAAAAA 3168

Qy   5073 AAAAAAAAAA 5082

Db   3169 AACCAAAAAA 3178
```

RESULT 12
AAN91578
ID AAN91578 standard; DNA; 3217 BP.
XX
AC AAN91578;
XX
DT 28-FEB-1990 (first entry)
XX
DE Rat androgen receptor DNA clone.
XX
KW Androgen receptor; TR2 polypeptide;
XX
OS Rat.
XX

FH	Key	Location/Qualifiers
FT	CDS	33..2739
FT		/*tag= a
FT		/product=98 kD polypeptide
FT	CDS	541..2739
FT		/*tag= b
FT		/product=79 kD polypeptide

XX
PN WO8909223-A.
XX
PD 05-OCT-1989.
XX
PF 24-MAR-1989; 89WO-US01238.
XX
PR 30-MAR-1988; 88US-0176107.
XX
PA (ARCH-) ARCH DEVELOPMENT CORP.
XX
PI Liao S, Chang C;
XX
DR WPI; 1989-309501/42.
DR P-PSDB; AAP91006.
XX
PT New DNA encoding new androgen receptor and TR2 polypeptide(s) - able
PT to bind DNA, and derived antibodies, useful for receptor assay and
PT purification.
XX
PS Claim 8; fig. 3; 60pp; English.
XX
CC The sequence is used to express the corresp. peptides and for
CC hybridisation assays of RNA and DNA encoding androgen receptors.
CC The 98 kD product starts at the first Met codon; the 79 kD product
CC starts from the second.
XX
SQ Sequence 3217 BP; 776 A; 874 C; 837 G; 730 T; 0 other;

Query Match 43.0%; Score 2185.8; DB 10; Length 3217;
Best Local Similarity 82.1%; Pred. No. 0;
Matches 2715; Conservative 0; Mismatches 407; Indels 185; Gaps 10;

```
Qy   1832 GAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCAGTTAGGGCTGGGAAGGG 1891

Db      1 GAATTCGGTGGAAGCTAGAGACAAGCTAAAGGATGGAGGTGCATTTAGGGCTGGGAAGGG 60

Qy   1892 TCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCAGAATCTGTTCCAGAGCG 1951
```

```
Qy   4061 ACGTGGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCA 4120
          | ||||| |||||||||||| |||||||||| |||||||||| || ||||| |||||||
Db   2167 ATGTGGATGACCAGATGGCAGTCATTCAGTATTCCTGGATGGGACTGATGGTATTTGCCA 2226

Qy   4121 TGGGCTGGCGATCCTTCACCAATGTCAACTCCAGGATGCTCTACTTCGCCCCTGATCTGG 4180
          |||| ||||| |||||||| |||||||||| |||||||||||||| || ||||| ||||
Db   2227 TGGGTTGGCGGTCCTTCACTAATGTCAACTCTAGGATGCTCTACTTTGCACCTGACCTGG 2286

Qy   4181 TTTTCAATGAGTACCGCATGCACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGC 4240
          |||||||||||| |||||||||||| || |||||||||||||| ||  | |||||||
Db   2287 TTTTCAATGAGTATCGCATGCACAAGTCTCGAATGTACAGCCAGTGCGTGAGGATGAGGC 2346

Qy   4241 ACCTCTCTCAAGAGTTTGGATGGCTCCAAATCACCCCCCAGGAATTCCTGTGCATGAAAG 4300
          |||| |||||||||||||||||||||| || ||||||||||||||||||||||||||||
Db   2347 ACCTTTCTCAAGAGTTTGGATGGCTCCAGATAACCCCCCAGGAATTCCTGTGCATGAAAG 2406

Qy   4301 CACTGCTACTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATG 4360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2407 CACTGCTACTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATG 2466

Qy   4361 AACTTCGAATGAACTACATCAAGGAACTCGATCGTATCATTGCATGCAAAAGAAAAAATC 4420
          |||||||||||||||||||||||||| ||||| ||||||||||||||||||||||||||
Db   2467 AACTTCGAATGAACTACATCAAGGAACTTGATCGCATCATTGCATGCAAAAGAAAAAATC 2526

Qy   4421 CCACATCCTGCTCAAGACGCTTCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTA 4480
          |||||||||||||||| |||||||||||||||||||||||||| || ||||||||||
Db   2527 CCACATCCTGCTCAAGGCGCTTCTACCAGCTCACCAAGCTCCTGGATTCTGTGCAGCCTA 2586

Qy   4481 TTGCGAGAGAGCTGCATCAGTTCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCG 4540
          |||| |||||||||||||| |||||||||||||||||||||||||| || ||||||||||
Db   2587 TTGCAAGAGAGCTGCATCAATTCACTTTTGACCTGCTAATCAAGTCCCATATGGTGAGCG 2646

Qy   4541 TGGACTTTCCGGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTG 4600
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2647 TGGACTTTCCTGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTG 2706

Qy   4601 GGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCAC 4660
          |||||||||||||||||||||||||||||| ||||||||  |||||||||||| | ||| |
Db   2707 GGAAAGTCAAGCCCATCTATTTCCACACACAGTGAAGATTTGGAAACCCTAATACCCAAA 2766

Qy   4661 CCCAGCTCATGCCCCCCTTTCAGATGTCTTCTGCCTGT--TATAACTCTGCACTACTCCTC 4718
          |||| ||  |  ||| |||||||||||||||||||  |||||||||||||||| |||
Db   2767 CCCACCTTGT-TCCCTTTTCAGATGTCTTCTGCCTGTTATATAACTCTGCACTACTTCTC 2825

Qy   4719 TGCAGTGCCTT-GGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAACATGTTCCTGA 4777
          |||||||||| |||||| |||||||| |||||||||||||||||||||||||||  |
Db   2826 TGCAGTGCCTTGGGGGAAATTCCTCTACTGATGTACAGTCTGTCATGAACATGTTCCCCA 2885

Qy   4778 ATTCTATTTGCTGGGCTTTTTTTTTCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCT 4837
          |||||||| ||||||||||             || |||||||||||||||||| || |||
Db   2886 GTTCTATTTCCTGGGCTTTT------------CCTTCTTTCTTTTTCTTCTTCTCTGCCT 2933

Qy   4838 ATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCTATCTGTGTTT 4897
          | | |||||||||||||| ||  || |  ||| |   |||| ||||||  ||||||||
Db   2934 CTTTTACCCTCCCATGGCACATTTTGAATCCGCTGCGTGTTGT-GCTCCTGCCTGTGTTT 2992

Qy   4898 TGAATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTG 4957
          |||| |||||||| ||| || |||||||||||| || | ||||||||||||| ||
Db   2993 TGAGTTTTGTTGTATTTCTTCAAGTCTGTGATGATCTTCTTGTGGCCCAGTGTCAA-CTG 3051

Qy   4958 TGCTTGTTTACAGCAC--TACTCTGTGCCAGCCACACAAACGTTTACTTATCTTATGCCA 5015
          |||||||||| |||||  | || ||||||| ||| |||| ||||||| | |||||||||
Db   3052 TGCTTGTTTATAGCACTGTGCTGTGTGCCAACCAAGCAAATGTTTACTCACCTTATGCCA 3111

Qy   5016 CGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAATCAAAACAAAAAACAAGCAAACAAA 5075
          || |||||||||||||||  | ||||| |||||  | ||  | | | |   ||| ||
Db   3112 TGGCAAGTTTAGAGAGCTATAAGTATCTTGGGAAGAAACAAACAGAGAGAGTAAAAAAAC 3171

Qy   5076 AAAAAAA 5082
          |||||||
Db   3172 CAAAAAA 3178
```

RESULT 13
ABK34012
ID ABK34012 standard; DNA; 3715 BP.
XX
AC ABK34012;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human DNA for staging of Astrocytomas #50.
XX
KW Human; ds; astrocytoma; cytostatic; staging; cysteine methylation; CpG;
KW bisulphite; brain tissue; MALDI; ESI; electron spray mass spectrometry;
KW matrix assisted laser desorption/ionization mass spectrometry.
XX
OS Homo sapiens.
XX
PN WO200202808-A2.
XX
PD 10-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07538.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-171649/22.
XX
PT Novel chemically modified genomic DNA sequences, useful in the
PT characterisation, classification, differentiation, grading, staging;
PT treatment and/or diagnosis of astrocytomas or predisposition to
PT astrocytomas -
XX
PS Claim 1; SEQ ID No 99; 37pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence (I) of at
CC least 18 bases in length of a segment of chemically pre-treated genomic
CC DNA which has any one of the sequences of (ABK33919-ABK34032) or its
CC complement. Also included are an oligonucleotide or peptide nucleic
CC acid (or set thereof) of at least 9 nucleotides which hybridises to (I),
CC primers for (I), probes for detecting cytosine methylation or single-
CC nucleotide polymorphisms (SNP) in (I), an array of oligomers
CC or peptide nucleic acids for analysing diseases associated with the
CC methylation states of the CpG dinucleotides of (I). The array is useful
CC for determining genetic and/or epigenetic parameters, classification,
CC differentiation, grading, staging, treatment and/or diagnosis of
CC astrocytomas, or the predisposition to astrocytomas by analysing cytosine
CC methylations, involves obtaining a biological sample containing genomic
CC DNA, extracting the genomic DNA, converting cytosine bases which are
CC unmethylated at the 5-position, in the genomic DNA sample, to uracil or
CC another base which is dissimilar to cytosine in terms of hybridisation
CC behaviour, by chemical treatment and amplifying chemically pre-treated
CC genomic DNA fragments using the array and a polymerase, where the
CC amplificates carry a detectable label. The method further involves
CC identifying methylation status of one or more cytosine positions, and
CC analysing methylation status of the cytosine positions by reference to
CC one or more data sets. The genomic DNA is chemically treated by using a
CC bisulphite, hydrogen sulphite or disulphite. The amplification
CC step amplifies DNA which is of particular interest in astrocytoma or
CC brain tissue, based on the specific genomic methylation status of brain
CC tissues, as opposed to background DNA. The amplificates carry a
CC fluorescent label or radionuclide. Optionally, the labels of the
CC amplificates are detachable molecule fragments having a typical mass
CC which are detected in a mass spectrometer. The fragments of chemically
CC pre-treated genomic DNA to be amplified, have a single positive or
CC negative charge for a better detectability in the mass spectrometer.
CC Preferably, the amplificates or fragments of the amplificates are
CC detected by matrix assisted laser desorption/ionization mass spectrometry
CC (MALDI) or using electron spray mass spectrometry (ESI). The

```
CC   present sequence is one of the chemically pre-treated reference DNA
CC   samples of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3715 BP; 841 A; 149 C; 1001 G; 1724 T; 0 other;

  Query Match             42.4%;  Score 2154;  DB 24;  Length 3715;
  Best Local Similarity   74.5%;  Pred. No. 0;
  Matches 2782;  Conservative    0;  Mismatches  905;  Indels   45;  Gaps    4;

Qy   1340 GCGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCCCACCC 1399
Db      9 GCGGAGAGAATTTTTTGTTTTTTTTTATTTTTTTTTTATTTTTTTTGTTTTTTTTATTT 68

Qy   1400 CGAGTGCGGAGCCAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCAAAAAA 1459
Db     69 CGAGTGCGGAG-TAGAGATTAAAAGATGAAAAGGTAGTTAGGTTTTTAGTAGTTAAAAAA 127

Qy   1460 CAAAACAAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTCAGTTCTTAT 1519
Db    128 TAAAATAAATAAAAATAAAAAAGTCGAAATAAAAGAAAAAGATAATAATTTAGTTTTTAT 187

Qy   1520 TTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTCTTTTA 1579
Db    188 TTGTATTTATTTTAGTGGATATTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTTTTTTA 247

Qy   1580 AGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGCCTAGC 1639
Db    248 AGATTTGGGTATTTTTTGAATTTATTTTTTAAGTATTAAGAGATAGATTGTGAGTTTAGT 307

Qy   1640 AGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGGCTGTCAGAGCG 1699
Db    308 AGGGTAGATTTTGTTTATCGTGTGTTTTTTTTTGTACGAGATTTTGAGGTTGTTAGAGCG 367

Qy   1700 CTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGGCAGCT 1759
Db    368 TTTTTTGCGTGGTTGTTTTCGTAAGTTTTTTTTTTTGGAGTTTTTCGTAGGTGGGTAGTT 427

Qy   1760 AGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAAGGGGA 1819
Db    428 AGTTGTAGCGATTATCGTATTATTATAGTTTGTTGAATTTTTTTGAGTAAGAGAAGGGGA 487

Qy   1820 GGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCAGTTAG 1879
Db    488 GGCGGGGTAAGGGAAGTAGGTGGAAGATTTAGTTAAGTTTAAGGATGGAAGTGTAGTTAG 547

Qy   1880 GGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCAGAATC 1939
Db    548 GGTTGGGAAGGGTTTATTTTCGGTCGTCGTTTAAGATTTATCGAGGAGTTTTTTAGAATT 607

Qy   1940 TGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGCCGCGA 1999
Db    608 TGTTTTAGAGCGTGCGCGAAGTGATTTAGAATTCGGGTTTTAGGTATTTAGAGGTCGCGA 667

Qy   2000 GCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGCAGCAGCAGCAGCAGC 2059
Db    668 GCGTAGTATTTTTCGGCGTTAGTTTGTTGTTGTT----------------------GT 703

Qy   2060 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGACTAGCC 2119
Db    704 AGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAAGAGATTAGTT 763

Qy   2120 CCAGGCAGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAGAGGCC 2179
Db    764 TTAGGTAGTAGTAGTAGTAGTAGGGTGAGGATGGTTTTTTTTAAGTTTATCGTAGAGGTT 823

Qy   2180 CCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGCCCTGG 2239
Db    824 TTATAGGTTATTTGGTTTTGGATGAGGAATAGTAATTTTTATAGTCGTAGTCGGTTTTGG 883

Qy   2240 AGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAGCAAGG 2299
Db    884 AGTGTTATTTCGAGAGAGGTTGCGTTTTAGAGTTTGGAGTCGTCGTGGTCGTTAGTAAGG 943

Qy   2300 GGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATCCACGT 2359
Db    944 GGTTGTCGTAGTAGTTGTTAGTATTTTCGGACGAGGATGATTTAGTTGTTTTATTTACGT 1003

Qy   2360 TGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAAAGACA 2419
Db   1004 TGTTTTTGTTGGGTTTTATTTTTTTCGGTTTAAGTAGTTGTTTCGTTGATTTTAAAGATA 1063

Qy   2420 TCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGTATCCG 2479
Db   1064 TTTTGAGCGAGGTTAGTATTATGTAATTTTTTTAGTAATAGTAGTAGGAAGTAGTATTCG 1123

Qy   2480 AAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAAGGACA 2539
Db   1124 AAGGTAGTAGTAGCGGGAGAGCGAGGGAGGTTTCGGGGGTTTTTATTTTTTTTAAGGATA 1183

Qy   2540 ATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGCAGTGT 2599
Db   1184 ATTATTTAGGGGGTATTTCGATTATTTTTGATAACGTTAAGGAGTTGTGTAAGGTAGTGT 1243

Qy   2600 CGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACAGCTTC 2659
Db   1244 CGGTGTTTATGGGTTTGGGTGTGGAGGCGTTGGAGTATTTGAGTTTAGGGGAATAGTTTC 1303

Qy   2660 GGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCCTT 2719
Db   1304 GGGGGGATTGTATGTACGTTTTATTTTTGGGAGTTTTATTCGTTGTGCGTTTTATTTTTT 1363

Qy   2720 GTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAGCACTG 2779
Db   1364 GTGTTTTATTGGTCGAATGTAAAGGTTTTTGTTAGACGATAGCGTAGGTAAGAGTATTG 1423

Qy   2780 AAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGGCGAGA 2839
Db   1424 AAGATATTGTTGAGTATTTTTTTTTTAAGGGAGGTTATATTAAAGGGTTAGAAGGCGAGA 1483

Qy   2840 GCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCCGTCTA 2899
Db   1484 GTTTAGGTTGTTTTGGTAGCGTTGTAGTAGGGAGTTTCGGGATATTTGAATTGTCGTTTA 1543

Qy   2900 CCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACT 2959
Db   1544 TTTTGTTTTTTTATAAGTTCGGAGTATTGGACGAGGTAGTTGCGTATTAGAGTCGCGATT 1603

Qy   2960 ACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCC 3019
Db   1604 ATTATAATTTTTTATTGGTTTTGGTCGGATCGTCGTTTTTTCGTCGTTTTTTTATTTTT 1663

Qy   3020 ACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGCGGCGG 3079
Db   1664 ACGTTCGTATTAAGTTGGAGAATTCGTTGGATTACGGTAGCGTTTGGGCGGTTGCGGCGG 1723

Qy   3080 CGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTT 3139
Db   1724 CGTAGTGTCGTTATGGGGATTTGGCGAGTTTGTATGGCGCGGGTGTAGCGGGATTCGGTT 1783

Qy   3140 CTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAG 3199
Db   1784 TTGGGTTATTTTTAGTCGTCGTTTTTTTATTTTGGTATATTTTTTTTATAGTCGAAGAAG 1843

Qy   3200 GCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGT---------GGCGGCGGCGGCG 3250
Db   1844 GTTAGTTGTATGGATCGTGTGGTGGTGGTGGGGGTGGTGGTGGCGGCGGCGGCGGCGCG 1903

Qy   3251 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAGCTGTAGCCC 3310
Db   1904 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGAAGTTGTAGTTT 1963

Qy   3311 CCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCACCGCAC 3370
```

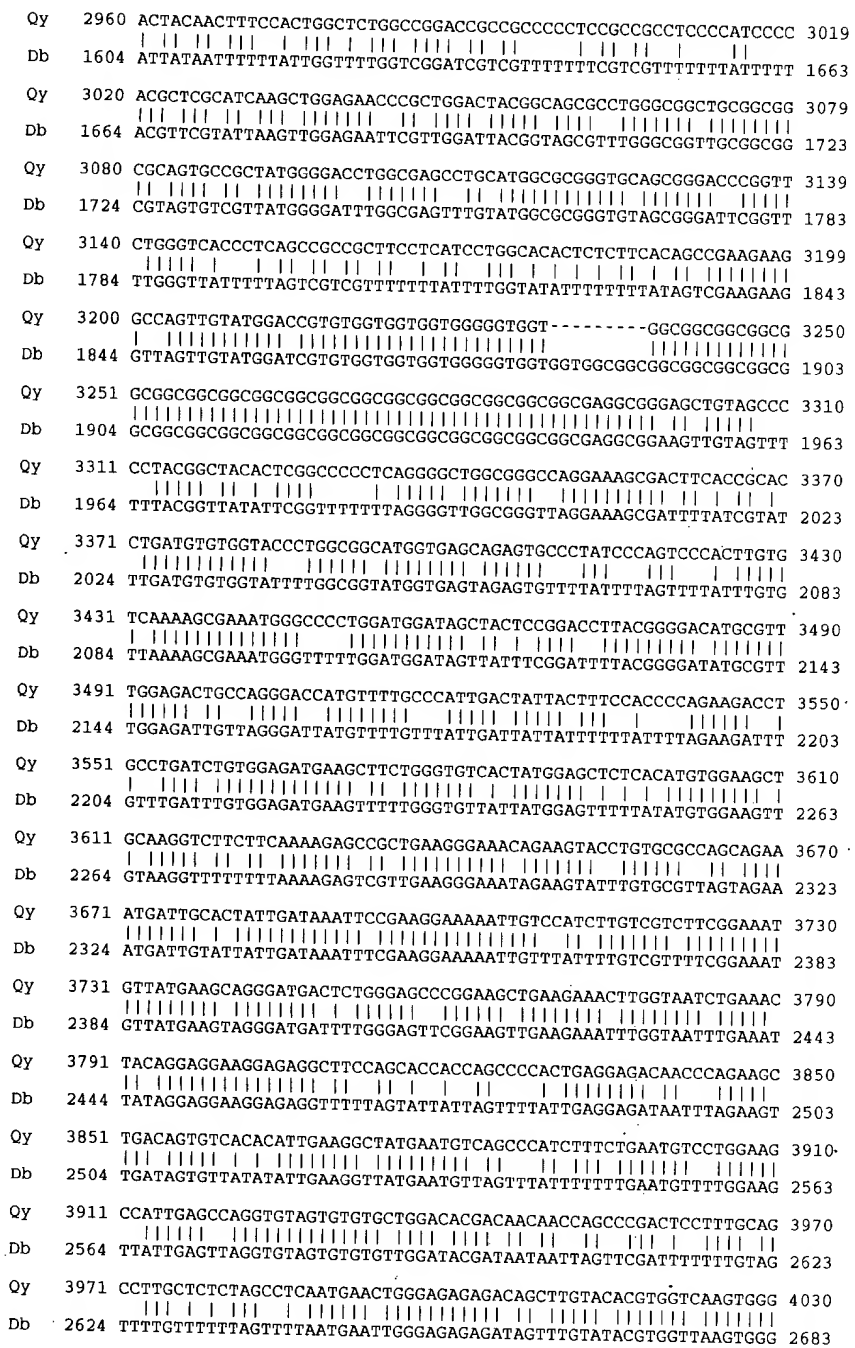
```
Db  1964 TTTACGGTTATATTCGGTTTTTTTAGGGGTTGGCGGGTTAGGAAAGCGATTTTATCGTAT 2023
Qy  3371 CTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCACTTGTG 3430
Db  2024 TTGATGTGTGGTATTTTGGCGGTATGGTGAGTAGAGTGTTTTATTTTAGTTTTATTTGTG 2083
Qy  3431 TCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACATGCGTT 3490
Db  2084 TTAAAAGCGAAATGGGTTTTTGGATGGATAGTTATTTCGGATTTTACGGGGATATGCGTT 2143
Qy  3491 TGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAAGACCT 3550
Db  2144 TGGAGATTGTTAGGGATTATGTTTTGTTTATTGATTATTATTTTTTTATTTTAGAAGATTT 2203
Qy  3551 GCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACATGTGGAAGCT 3610
Db  2204 GTTTGATTTGTGGAGATGAAGTTTTTGGGTGTTATTATGGAGTTTTTATATGTGGAAGTT 2263
Qy  3611 GCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAGCAGAA 3670
Db  2264 GTAAGGTTTTTTTTAAAAGAGTCGTTGAAGGGAAATAGAAGTATTTGTGCGTTAGTAGAA 2323
Qy  3671 ATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCGGAAAT 3730
Db  2324 ATGATTGTATTATTGATAAATTTCGAAGGAAAAATTGTTTATTTTGTCGTTTTCGGAAAT 2383
Qy  3731 GTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCTGAAAC 3790
Db  2384 GTTATGAAGTAGGGATGATTTTGGGAGTTCGGAAGTTGAAGAAATTTGGTAATTTGAAAT 2443
Qy  3791 TACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCAGAAGC 3850
Db  2444 TATAGGAGGAAGGAGAGGTTTTTAGTATTATTAGTTTTATTGAGGAGATAATTTAGAAGT 2503
Qy  3851 TGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATGTCCTGGAAG 3910
Db  2504 TGATAGTGTTATATATTGAAGGTTATGAATGTTAGTTTATTTTTTTGAATGTTTTGGAAG 2563
Qy  3911 CCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTTTGCAG 3970
Db  2564 TTATTGAGTTAGGTGTAGTGTGTGTTGGATACGATAATAATTAGTTCGATTTTTTTGTAG 2623
Qy  3971 CCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAAGTGGG 4030
Db  2624 TTTTGTTTTTTAGTTTTAATGAATTGGGAGAGAGATAGTTTGTATACGTGGTTAAGTGGG 2683
Qy  4031 CCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCATTCAGT 4090
Db  2684 TTAAGGTTTTGTTTGGTTTTCGTAATTTATACGTGGACGATTAGATGGTTGTTATTTAGT 2743
Qy  4091 ACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGTCAACT 4150
Db  2744 ATTTTTGGATGGGGTTTATGGTGTTTGTTATGGGTTGGCGATTTTTTATTAATGTTAATT 2803
Qy  4151 CCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCC 4210
Db  2804 TTAGGATGTTTTATTTCGTTTTTGATTTGGTTTTTAATGAGTATCGTATGTATAAGTTTC 2863
Qy  4211 GGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCAAA 4270
Db  2864 GGATGTATAGTTAGTGTGTTCGAATGAGGTATTTTTTTTAAGAGTTTGGATGGTTTTAAA 2923
Qy  4271 TCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCCAGTGG 4330
Db  2924 TTATTTTTTAGGAATTTTTGTGTATGAAAGTATTGTTATTTTTTAGTATTATTTTAGTGG 2983
Qy  4331 ATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTCG 4390
Db  2984 ATGGGTTGAAAAATTAAAAATTTTTTGATGAATTTCGAATGAATTATATTAAGGAATTCG 3043
Qy  4391 ATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTACCAGC 4450
Db  3044 ATCGTATTATTGTATGTAAAAGAAAAAATTTTATATTTTGTTTAAGACGTTTTTATTAGT 3103
Qy  4451 TCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCACTTTTG 4510
Db  3104 TTATTAAGTTTTTGGATTTCGTGTAGTTTATTGCGAGAGAGTTGTATTAGTTTATTTTTG 3163
Qy  4511 ACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGAGATCA 4570
Db  3164 ATTTGTTAATTAAGTTATATATGGTGAGCGTGGATTTTTCGGAAATGATGGTAGAGATTA 3223
Qy  4571 TCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACCC 4630
Db  3224 TTTTTGTGTAAGTGTTTAAGATTTTTTTTGGGAAAGTTAAGTTTATTTATTTTTATATTT 3283
Qy  4631 AGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATGTCTTC 4690
Db  3284 AGTGAAGTATTGGAAATTTTATTTTTTTATTTTAGTTTATGTTTTTTTTTAGATGTTTTT 3343
Qy  4691 TGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTATTGATG 4750
Db  3344 TGTTTGTTATAATTTTGTATTATTTTTTTGTAGTGTTTTGTTTAATTTTTTTTATTGATG 3403
Qy  4751 TACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTCTCTTTCT 4810
Db  3404 TATAGTTTGTTATG----------GAATTTTATTTGTTGGGTTTTTTTTTTTTTTTTTT 3452
Qy  4811 CTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTTGC 4870
Db  3453 TTTTTTTTTTTTTTTTTTTTTTTTTTTATTTAATTTTTTTATGGTATTTTTAGATTTTGT 3512
Qy  4871 TTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATG 4930
Db  3513 TTTTTATTGTGGTTTTTATTTGTGTTTTGAATGGTGTTGTATGTTTTTAAATTTGTGATG 3572
Qy  4931 ATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCAGCCAC 4990
Db  3573 ATTTTTATATGGTTTAGTGTTAAGTTGTGTTTGTTTATAGTATTATTTTGTGTTAGTTAT 3632
Qy  4991 ACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAA 5050
Db  3633 ATAAACGTTTATTTATTTTATGTTACGGGAAGTTTAGAGAGTTAAGATTATTTGCGGAAA 3692
Qy  5051 TCAAAACAAAAA 5062
Db  3693 TTAAAATAAAAA 3704
```

RESULT 14
ABK31544
ID ABK31544 standard; DNA; 3715 BP.
XX
AC ABK31544;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified DNA #194.
XX
KW Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200926-A2.
XX
PD 03-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP07472.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX

PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-147896/19.
XX
PT Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction -
XX
PS Claim 1; SEQ ID No 387; 24pp; English.
XX
CC The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or
CC disulphite. Also disclosed are oligonucleotides and/or PNA oligomers
CC for detecting the cytosine methylation state (CpG islands) of these
CC genes, and a method for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with signal transduction.
CC The genomic DNA can be obtained from cells or cellular components which
CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
CC histologic object slides, and all their possible combinations. The
CC sequences of the invention are useful for the diagnosis and therapy of
CC diseases associated with signal transduction e.g. solid tumours and
CC cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA
CC sequences of different genes associated with signal transduction, or
CC their complementary sequences.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 3715 BP; 841 A; 149 C; 1001 G; 1724 T; 0 other;

Query Match 42.4%; Score 2154; DB 24; Length 3715;
Best Local Similarity 74.5%; Pred. No. 0;
Matches 2782; Conservative 0; Mismatches 905; Indels 45; Gaps 4;

```
Qy   1340 GCGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCTCCTCCTGCCTTCCCCACCC 1399
          ||||||||||   | ||||||     | | | | |  | |  ||  ||    |
Db      9 GCGGAGAGAATTTTTTGTTTTTTTTTTATTTTTTTTTTATTTTTTTTTGTTTTTTTTATTT 68

Qy   1400 CGAGTGCGGAGCCAGAGATCAAAAGATGAAAAGGCAGTCAGGTCTTCAGTAGCCAAAAAA 1459
          ||||||||||||  ||||||  ||||||||||||||| ||| |||| || |||||  ||||||
Db     69 CGAGTGCGGAG-TAGAGATTAAAAGATGAAAAGGTAGTTAGGTTTTTAGTAGTTAAAAAA 127

Qy   1460 CAAAACAAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAGATAATAACTGAGTTCTTAT 1519
          ||||  |||  |||||  |||||||  |||||||||||||||||||||||||  | ||||  ||||
Db    128 TAAAATAAATAAAAATAAAAAAGTCGAAATAAAAGAAAAAGATAATAATTTAGTTTTTAT 187

Qy   1520 TTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTCTTTTA 1579
          |||  |  || ||  ||||||  | |||||||||||||||||||||||||||||||  |||||
Db    188 TTGTATTTATTTTAGTGGATATTGAATTTGGAAGGTGGAGGATTTTGTTTTTTTTTTTTA 247

Qy   1580 AGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGAGACAGACTGTGAGCCTAGC 1639
          ||||  ||||  ||  ||||||||  ||   ||  ||||||||||||||  |||  ||||||  |||
Db    248 AGATTTGGGTATTTTTTGAATTTATTTTTTAAGTATTAAGAGATAGATTGTGAGTTTAGT 307

Qy   1640 AGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGACTTTGAGGCTGTCAGAGCG 1699
          ||||  ||||  ||||  | |||||||  ||  ||  ||  ||||||  |||||||  |||  ||||||
Db    308 AGGGTAGATTTTGTTTATCGTGTGTTTTTTTTTGTACGAGATTTTGAGGTTGTTAGAGCG 367

Qy   1700 CTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGGCAGCT 1759
          ||||||||||||||  |  ||  ||||||  ||  |  |||||  ||  ||  |||||||  ||  |
Db    368 TTTTTTGCGTGGTTGTTTTCGTAAGTTTTTTTTTTTGGAGTTTTTCGTAGGTGGGTAGTT 427

Qy   1760 AGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAAGGGGA 1819
          ||  ||  |||||  ||  ||  ||  |  ||  |||||||  |  ||  ||||||||||||
Db    428 AGTTGTAGCGATTATCGTATTATTATAGTTTGTTGAATTTTTTTGAGTAAGAGAAGGGGA 487

Qy   1820 GGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCAGTTAG 1879
          |||||||||||||||||||||||||||||  ||  |||  | ||||||||||||||  ||||||

Db    488 GGCGGGGTAAGGGAAGTAGGTGGAAGATTTAGTTAAGTTTAAGGATGGAAGTGTAGTTAG 547

Qy   1880 GGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCAGAATC 1939
          ||  ||||||||||  ||   ||||  ||  |||  ||||  ||  |||||||  |||  |||||
Db    548 GGTTGGGAAGGGTTTATTTTCGGTCGTCGTTTAAGATTTATCGAGGAGTTTTTTAGAATT 607

Qy   1940 TGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGCCGCGA 1999
          ||||  ||||||||||||||||||  ||||  ||||    |||  |   |||||  |||||
Db    608 TGTTTTAGAGCGTGCGCGAAGTGATTTAGAATTCGGGTTTTAGGTATTTAGAGGTCGCGA 667

Qy   2000 GCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGCAGCAGGAGCAGCAGC 2059
          |||  ||  |   |  |||||  ||||||  ||  ||  |                    |
Db    668 GCGTAGTATTTTTCGGCGTTAGTTTGTTGTTGTT----------GT 703

Qy   2060 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGACTAGCC 2119
          ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||||||  |||
Db    704 AGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAAGAGATTAGTT 763

Qy   2120 CCAGGCAGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAGAGGCC 2179
            |||  ||  ||  ||  ||  ||  ||||||||||||||||  |    |||  ||||||||||
Db    764 TTAGGTAGTAGTAGTAGTAGTAGGGTGAGGATGGTTTTTTTTAAGTTTATCGTAGAGGTT 823

Qy   2180 CCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGCCCTGG 2239
            |  |||  ||  ||||  ||||||||||||  ||  ||  ||  |  ||  ||  ||||||  |||
Db    824 TTATAGGTTATTTGGTTTTGGATGAGGAATAGTAATTTTTATAGTCGTAGTCGGTTTTGG 883

Qy   2240 AGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAGCAAGG 2299
          ||||  |  ||||||||||||||||  ||||  |||||  ||  |||||  ||||
Db    884 AGTGTTATTTCGAGAGAGGTTGCGTTTTAGAGTTTGGAGTCGTCGTGGTCGTTAGTAAGG 943

Qy   2300 GGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATCCACGT 2359
          ||  ||  ||  ||  ||  ||  ||  |  |  ||||||||||||||  |  ||  ||  ||||
Db    944 GGTTGTCGTAGTAGTTGTTAGTATTTTCGGACGAGGATGATTTAGTTGTTTTTATTTACGT 1003

Qy   2360 TGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAAAGACA 2419
          |||  ||  ||||  |  |||  |||  |||||  ||  ||  |  ||  |||  |||||||  |
Db    1004 TGTTTTTGTTGGGTTTTATTTTTTTCGGTTTAAGTAGTTGTTTCGTTGATTTTAAAGATA 1063

Qy   2420 TCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGTATCCG 2479
          |  |||||||||  ||  |  |||  ||  |  ||  ||  ||  ||  ||||||  |||||  ||
Db    1064 TTTTGAGCGAGGTTAGTATTATGTAATTTTTTTAGTAATAGTAGTAGGAAGTAGTATTCG 1123

Qy   2480 AAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAAGGACA 2539
          ||||  ||  ||  ||||||||||||||||||||  |||||||  |  ||  |  |||||  |
Db    1124 AAGGTAGTAGTAGCGGGAGAGCGAGGGAGGTTTCGGGGGTTTTTATTTTTTTTAAGGATA 1183

Qy   2540 ATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGCAGTGT 2599
          ||||  ||||||||  |  |||||  ||||  |||  ||||  ||||||||||||||||  |||||
Db    1184 ATTATTTAGGGGGTATTTCGATTATTTTTGATAACGTTAAGGAGTTGTGTAAGGTAGTGT 1243

Qy   2600 CGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACAGCTTC 2659
          ||||||  |||||  ||||||||||||||||||||  ||  |||||  |||||||  ||  |||
Db    1244 CGGTGTTTATGGGTTTGGGTGTGGAGGCGTTGGAGTATTTGAGTTTAGGGGAATAGTTTC 1303

Qy   2660 GGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCCTT 2719
          ||||||||||  |||||||    |  ||||||||||  |  ||  |||||||  |  ||
Db    1304 GGGGGGATTGTATGTACGTTTTATTTTTGGGAGTTTTATTCGTTGTGCGTTTTATTTTTT 1363

Qy   2720 GTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAGCACTG 2779
          |||  |||||  ||||||  |  ||  |||||||  ||||  |||  |||||  |  ||
Db    1364 GTGTTTTATTGGTCGAATGTAAAGGTTTTTTGTTAGACGATAGCGTAGGTAAGAGTATTG 1423

Qy   2780 AAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGGCGAGA 2839
          ||||||  ||  ||||||||  |||  ||||||||||  |  ||||||  |||||||||||||
Db    1424 AAGATATTGTTGAGTATTTTTTTTTTAAGGGAGGTTATATTAAAGGGTTAGAAGGCGAGA 1483

Qy   2840 GCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCCGTCTA 2899
          |  ||||  ||  |  |||  ||||  ||  ||  ||||||  |  |||||  ||  ||  ||  ||
Db    1484 GTTTAGGTTGTTTTGGTAGCGTTGTAGTAGGGAGTTTCGGGATATTTGAATTGTCGTTTA 1543

Qy   2900 CCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACT 2959
          |||  |  |  ||  ||||  |||||  |  |||||||||  ||  ||||||  ||||||||  |  |
Db    1544 TTTTGTTTTTTTATAAGTTCGGAGTATTGGACGAGGTAGTTGCGTATTACAGTCGCGATT 1603
```


Qy 2960 ACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCC 3019
Db 1604 ATTATAATTTTTTATTGGTTTTGGTCGGATCGTCGTTTTTTTCGTCGTTTTTTTATTTTT 1663
Qy 3020 ACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGCGGCGG 3079
Db 1664 ACGTTCGTATTAAGTTGGAGAATTCGTTGGATTACGGTAGCGTTTGGGCGGTTGCGGCGG 1723
Qy 3080 CGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTT 3139
Db 1724 CGTAGTGTCGTTATGGGGATTTGGCGAGTTTGTATGGCGCGGGTGTAGCGGGATTCGGTT 1783
Qy 3140 CTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAG 3199
Db 1784 TTGGGTTATTTTTAGTCGTCGTTTTTTTATTTTGGTATATTTTTTTTATAGTCGAAGAAG 1843
Qy 3200 GCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGT---------GGCGGCGGCGGCG 3250
Db 1844 GTTAGTTGTATGGATCGTGTGGTGGTGGTGGGGGTGGTGGTGGCGGCGGCGGCGGCG 1903
Qy 3251 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAGCTGTAGCCC 3310
Db 1904 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGAAGTTGTAGTTT 1963
Qy 3311 CCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCACCGCAC 3370
Db 1964 TTTACGGTTATATTCGGTTTTTTTAGGGGTTGGCGGGTTAGGAAAGCGATTTTATCGTAT 2023
Qy 3371 CTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCACTTGTG 3430
Db 2024 TTGATGTGTGGTATTTTGGCGGTATGGTGAGTAGAGTGTTTTATTTTAGTTTTATTTGTG 2083
Qy 3431 TCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACATGCGTT 3490
Db 2084 TTAAAAGCGAAATGGGTTTTTGGATGGATAGTTATTTCGGATTTTACGGGGATATGCGTT 2143
Qy 3491 TGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAAGACCT 3550
Db 2144 TGGAGATTGTTAGGGATTATGTTTTGTTTATTGATTATTATTTTTTATTTTAGAAGATTT 2203
Qy 3551 GCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACATGTGGAAGCT 3610
Db 2204 GTTTGATTTGTGGAGATGAAGTTTTTGGGTGTTATTATGGAGTTTTTATATGTGGAAGTT 2263
Qy 3611 GCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAGCAGAA 3670
Db 2264 GTAAGGTTTTTTTTAAAAGAGTCGTTGAAGGGAAATAGAAGTATTTGTGCGTTAGTAGAA 2323
Qy 3671 ATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCGGAAAT 3730
Db 2324 ATGATTGTATTATTGATAAATTTCGAAGGAAAAATTGTTTATTTTGTCGTTTTCGGAAAT 2383
Qy 3731 GTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGTAATCTGAAAC 3790
Db 2384 GTTATGAAGTAGGGATGATTTTGGGAGTTCGGAAGTTGAAGAAATTTGGTAATTTGAAAT 2443
Qy 3791 TACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACAACCCAGAAGC 3850
Db 2444 TATAGGAGGAAGGAGAGGTTTTTAGTATTATTAGTTTTATTGAGGAGATAATTTAGAAGT 2503
Qy 3851 TGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAATGTCCTGGAAG 3910
Db 2504 TGATAGTGTTATATATTGAAGGTTATGAATGTTAGTTTATTTTTTTGAATGTTTTGGAAG 2563
Qy 3911 CCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGACTCCTTTGCAG 3970
Db 2564 TTATTGAGTTAGGTGTAGTGTGTGTTGGATACGATAATAATTAGTTCGATTTTTTTGTAG 2623
Qy 3971 CCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTGGTCAAGTGGG 4030
Db 2624 TTTTGTTTTTTAGTTTTAATGAATTGGGAGAGAGATAGTTTGTATACGTGGTTAAGTGGG 2683
Qy 4031 CCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCTGTCATTCAGT 4090
Db 2684 TTAAGGTTTTGTTTGGTTTTCGTAATTTATACGTGGACGATTAGATGGTTGTTATTTAGT 2743
Qy 4091 ACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGTCAACT 4150
Db 2744 ATTTTTGGATGGGGTTTATGGTGTTTGTTATGGGTTGGCGATTTTTTATTAATGTTAATT 2803
Qy 4151 CCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCC 4210
Db 2804 TTAGGATGTTTTATTTCGTTTTTGATTTGGTTTTTAATGAGTATCGTATGTATAAGTTTC 2863
Qy 4211 GGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCAAA 4270
Db 2864 GGATGTATAGTTAGTGTGTTCGAATGAGGTATTTTTTTTAAGAGTTTGGATGGTTTTAAA 2923
Qy 4271 TCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCCAGTGG 4330
Db 2924 TTATTTTTTAGGAATTTTTGTGTATGAAAGTATTGTTATTTTTTAGTATTATTTTAGTGG 2983
Qy 4331 ATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTCG 4390
Db 2984 ATGGGTTGAAAAATTAAAAATTTTTTGATGAATTTCGAATGAATTATATTAAGGAATTCG 3043
Qy 4391 ATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGCTTCTACCAGC 4450
Db 3044 ATCGTATTATTGTATGTAAAAGAAAAAATTTTATATTTTGTTTAAGACGTTTTTATTAGT 3103
Qy 4451 TCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGTTCACTTTTG 4510
Db 3104 TTATTAAGTTTTTGGATTTCGTGTAGTTTATTGCGAGAGAGTTGTATTAGTTTATTTTTG 3163
Qy 4511 ACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGAGATCA 4570
Db 3164 ATTTGTTAATTAAGTTATATATGGTGAGCGTGGATTTTTCGGAAATGATGGTAGAGATTA 3223
Qy 4571 TCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACCC 4630
Db 3224 TTTTTGTGTAAGTGTTTAAGATTTTTTTTGGGAAAGTTAAGTTTATTTATTTTTATATTT 3283
Qy 4631 AGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATGTCTTC 4690
Db 3284 AGTGAAGTATTGGAAATTTTATTTTTTTATTTTAGTTTATGTTTTTTTTAGATGTTTTT 3343
Qy 4691 TGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCCTCTATTGATG 4750
Db 3344 TGTTTGTTATAATTTTGTATTATTTTTTTGTAGTGTTTTGTTTAATTTTTTTTATTGATG 3403
Qy 4751 TACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTCTCTTTCT 4810
Db 3404 TATAGTTTGTTATG----------GAATTTTATTTGTTGGGTTTTTTTTCTTTTTTTT 3452
Qy 4811 CTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTTGC 4870
Db 3453 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTAATTTTTTTATGGTATTTTTAGATTTTGT 3512
Qy 4871 TTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATG 4930
Db 3513 TTTTTATTGTGGTTTTTATTTGTGTTTTGAATGGTGTTGTATGTTTTTAAATTTGTGATG 3572
Qy 4931 ATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCAGCCAC 4990
Db 3573 ATTTTTATATGGTTTAGTGTTAAGTTGTGTTTGTTTATAGTATTATTTTGTGTTAGTTAT 3632
Qy 4991 ACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAA 5050
Db 3633 ATAAACGTTTATTTATTTTATGTTACGGGAAGTTTAGAGAGTTAAGATTATCTTGGGGAAA 3692
Qy 5051 TCAAAACAAAAA 5062
Db 3693 TTAAAATAAAAA 3704

```
RESULT 15
AAF84342
ID   AAF84342 standard; DNA; 1810 BP.
XX
AC   AAF84342;
XX
DT   20-JUN-2001  (first entry)
XX
DE   Human androgen receptor gene.
XX
KW   Human; PCR primer; X chromosome inactivation; androgen receptor;
KW   methylation; HUMARA; ds.
XX
OS   Homo sapiens.
XX
PN   JP2001017199-A.
XX
PD   23-JAN-2001.
XX
PF   08-JUL-1999;   99JP-0194555.
XX
PR   08-JUL-1999;   99JP-0194555.
XX
PA   (MITP ) MITSUBISHI YUKA BCL KK.
XX
DR   WPI; 2001-285412/30.
XX
PT   Analysis of inactivated X chromosome useful for determining clonality
PT   and uniformity of growth of a cell population comprises analyzing
PT   methylation of human androgen receptor gene by a methyl-specific
PT   polymerase chain reaction -
XX
PS   Claim 2; Page 11-12; 14pp; Japanese.
XX
CC   The present invention relates to a method for analysis of inactivation of
CC   X chromosome. The method comprises analysing methylation of human
CC   androgen receptor (HUMARA) gene (the present sequence) by a
CC   methyl-specific polymerase chain reaction (PCR). The PCR primers amplify
CC   the base sequence of the region containing repeated number of
CC   polymorphism of CAG of the HUMARA gene, specific to methylation of the
CC   cytosine base at the 199th, the 203rd and the 206th or the 296th position
CC   of the present sequence. The method is useful for the detection of
CC   uniformity of cell growth, that is clonality.
XX
SQ   Sequence 1810 BP; 353 A; 565 C; 585 G; 307 T; 0 other;

  Query Match             33.8%;  Score 1715.6;  DB 22;  Length 1810;
  Best Local Similarity   98.7%;  Pred. No. 1.8e-284;
  Matches 1767;  Conservative    0;  Mismatches    4;  Indels   19;  Gaps    3;

Qy   1701 TTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGCTTCCCGCAGGTGGGCAGCTA 1760
           ||||||||||||||||||| |||||||||||||||||||||||| ||||||||||||
Db      1 TTTTTGCGTGGTTGCTCCCG-AAGTTTCCTTCTCTGGAGCTTCCCGCACGTGGGCAGCTA 59

Qy   1761 GCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAAGGGGAG 1820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 GCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCTTCTGAGCAAGAGAAGGGGAG 119

Qy   1821 GCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCAGTTAGG 1880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 GCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGGATGGAAGTGCAGTTAGG 179

Qy   1881 GCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCAGAATCT 1940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 GCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTACCGAGGAGCTTTCCAGAATCT 239

Qy   1941 GTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGCCGCGAG 2000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 GTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCCAGGCACCCAGAGGCCGCGAG 299

Qy   2001 CGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAGCAGCAGCAGCAGCAGCA 2060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 CGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCT----------GCAGCAGCAGCA 344

Qy   2061 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGACTAGCCC 2120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    345 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGACTAGCCC 404

Qy   2121 CAGGCAGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAGAGGCCC 2180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    405 CAGGCAGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAAGCCCATCGTAGAGGCCC 464

Qy   2181 CACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGCCCTGGA 2240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    465 CACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCACAGCCGCAGTCGGCCCTGGA 524

Qy   2241 GTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAGCAAGGG 2300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    525 GTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGCCAGCAAGGG 584

Qy   2301 GCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATCCACGTT 2360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    585 GCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGACTCAGCTGCCCCATCCACGTT 644

Qy   2361 GTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAAAGACAT 2420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    645 GTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCGCTGACCTTAAAGACAT 704

Qy   2421 CCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGTATCCGA 2480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    705 CCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCAGCAGGAAGCAGTATCCGA 764

Qy   2481 AGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCTCCCACTTCCTCCAAGGACAA 2540
           |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db    765 AGGCAGCAGCAGCGGGAGAGCGAGGGAGCGCTCGGGGGCTCCCACTTCCTCCAAGGACAA 824

Qy   2541 TTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGCAGTGTC 2600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    825 TTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAGGAGTTGTGTAAGGCAGTGTC 884

Qy   2601 GGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACAGCTTCG 2660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    885 GGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTGAGTCCAGGGGAACAGCTTCG 944

Qy   2661 GGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCCTTG 2720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    945 GGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCCTTG 1004

Qy   2721 TGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAGCACTGA 2780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1005 TGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGACAGCGCAGGCAAGAGCACTGA 1064

Qy   2781 AGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGGCGAGAG 2840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1065 AGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACCAAAGGGCTAGAAGGCGAGAG 1124

Qy   2841 CCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCCGTCTAC 2900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1125 CCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACACTTGAACTGCCGTCTAC 1184

Qy   2901 CCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACTA 2960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1185 CCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACTA 1244

Qy   2961 CTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCCA 3020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1245 CTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCCA 1304

Qy   3021 CGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGCGGCGGC 3080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1305 CGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGCGGCGGC 1364

Qy   3081 GCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTTC 3140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1365 GCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTTC 1424
```

```
Qy   3141 TGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAGG 3200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1425 TGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAGG 1484

Qy   3201 CCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGCGGCGGCGGCGGCGGCGGCGG 3260
          ||||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db   1485 CCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGT---GGCGGCGGCGGCGGCGGCGG 1541

Qy   3261 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAGCTGTAGCCCCCTACGGCTA 3320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1542 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGAGCTGTAGCCCCCTACGGCTA 1601

Qy   3321 CACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCACCGCACCTGATGTGTG 3380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1602 CACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCACCGCACCTGATGTGTG 1661

Qy   3381 GTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCACTTGTGTCAAAAGCGA 3440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1662 GTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCACTTGTGTCAAAAGCGA 1721

Qy   3441 AATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACATGCGTT 3490
          ||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   1722 AATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACATGCGGT 1771
```

Search completed: May 8, 2003, 15:16:10
Job time : 1176 secs